TWIN-ARGININE TRANSLOCASE MUTATIONS THAT SUPPRESS FOLDING QUALITY CONTROL AND PERMIT EXPORT OF MISFOLDED SUBSTRATE PROTEINS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Mark Alexander Rocco

May 2011

TWIN-ARGININE TRANSLOCASE MUTATIONS THAT SUPPRESS FOLDING QUALITY CONTROL AND PERMIT EXPORT OF MISFOLDED SUBSTRATE PROTEINS

Mark Alexander Rocco, Ph.D.

Cornell University 2011

The twin-arginine translocation (Tat) pathway is a robust protein translocation system capable of transporting cellular proteins across the bacterial cytoplasmic membrane. A hallmark of the bacterial Tat pathway is its ability to export proteins that have attained a fully-folded structure prior to export. An inbuilt feature of the Tat system is a folding quality control (QC) mechanism that discriminates between unfolded/misfolded and correctly folded proteins, allowing only the latter to pass through the inner membrane. This mechanism holds great promise for biotechnology applications because it has the potential to ensure the quality and structural integrity of exported proteins. Yet, despite the significance of the Tat system, the precise mechanism by which the TatABC protein complex delivers folded proteins into the periplasm remains poorly understood. Likewise, very little is known about the folding QC mechanism. Studies from our laboratory suggest that the QC mechanism resides within the TatABC proteins. Thus, to shed more light on this enigmatic process, we implemented a directed co-evolution strategy whereby all three components of the Tat pathway were evolved simultaneously to isolate suppressors that relaxed the QC feature. Selection of suppressors was enabled by a genetic selection comprised of a poorly folded substrate protein (α_3B) fused to the selectable reporter beta-lactamase (*Bla*). By screening recombinant libraries of the *tatABC* operon, we were able to

isolate variants that exported previously translocation-incompetent substrates and non-native Tat substrates. The resulting TatABC variants represent putative suppressors of the QC mechanism. Further information gained from these suppressors should provide a detailed understanding of the molecular mechanisms involved in TatABC-mediated QC. Moreover, these studies demonstrate that directed co-evolution of cellular protein machinery is a viable new strategy for dissecting complicated, poorly understood multiprotein complexes in living cells.

BIOGRAPHICAL SKETCH

Mark graduated from Northwestern University in Evanston, Illinois in June of 2006 with a Bachelor of Science in Materials Science & Engineering and a double major in Psychology. Immediately upon graduation Mark enrolled at Cornell University, where he would proceed to earn a Master of Science and Doctor of Philosophy in Biomedical Engineering. During his second year at Cornell University, Mark was awarded a Baden-Württemberg Stipendium from the German government funding a year-long research project in the laboratory of Prof. Dr. Bernd Bukau at the University of Heidelberg in Germany. Upon completion of his doctoral dissertation, Mark will pursue a brief post-doctoral position at the Garvan Institute of Medical Research in Sydney, Australia under the auspices of a National Science Foundation East Asia and Pacific Summer Institutes Fellowship. Thereafter, he will be funded by a Whitaker International Scholarship to conduct a second post-doctoral fellowship in the United Kingdom at Imperial College London.

This work is dedicated to my loving and caring family.

To my parents, Michaelangelo and Marilyn, without whom none of this would have been possible.

To my brother, Michaelangelo, for his guidance and wisdom.

And to my sister, Erica, for her endless support.

ACKNOWLEDGMENTS

This work was performed with support from the following funding sources (all awarded to Dr. Matthew P. DeLisa, unless noted otherwise):

James D. Watson Young Investigator Award presented by the New York State Office of Science, Technology, and Academic Research (NYSTAR)

Distinguished Faculty Award presented by the New York State Office of Science, Technology, and Academic Research (NYSTAR)

National Science Foundation (NSF) CAREER Award (BES #0449080)

Vybion, Inc.

Baden-Württemberg Stipendium (awarded to Dr. Mark Alexander Rocco)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION TO THE TWIN-ARGININE TRANSLOCATION (TAT) PATHWAY IN *ESCHERICHIA COLI* AND GENERAL QUALITY CONTROL MECHANISMS WITHIN LIVING CELLS

1.0 Introduction

In order for a protein to be functional, it must first fold properly within the cell. However, not all proteins fold the same way, i.e., proteins' native folds are remarkably variable. For example, some proteins fold very easily and quickly without the need for additional assistance, while others require the coordination of a cofactor (*1*) or binding partner (*2, 3*), and still others require post-translational modifications such as disulfide bond formation (*4*) or glycosylation, to gain protein function. A testament to the diversity and complexity of protein folding is the immense range of folding kinetics. It is generally recognized that the speed of protein folding can vary over eight orders of magnitude, with the fastest known folder able to achieve its native structure within just a few hundred nanoseconds (*5-7*). The birth of recombinant DNA technology has in many ways further exacerbated this fundamental protein folding problem in that many non-native, engineered proteins are required to fold properly in order to function in their new host, however, very little information can be predicted from their primary amino acid sequences alone. Moreover, an ongoing challenge to this field is interpreting how single mutations will, and do, affect protein folding (*5, 8*). Despite this perplexing conundrum, all organisms have evolved mechanisms to assist in the process of protein folding, from chaperones to proteolysis and almost everything in between.

The work presented herein will systematically address each of these aspects of protein folding quality control and proofreading, with an overall focus on the process of translocation of folded targets. We will follow the general path that a nascent polypeptide would take within the cell, that is, we will first address accessory factors and molecular chaperones that aid in protein folding, followed by the translocation processes which deliver preproteins to their final destinations, and finally we will examine several protein degradation pathways which exist in order to cleanup misfolded or otherwise unusable protein substrates.

1.1 General molecular chaperones that assist in the folding of nascent polypeptides

Quality control of the protein folding landscape begins with molecular chaperones within the cytoplasm of the cell and is a form of protein proofreading. Many molecular chaperones exist in organisms, from eukaryotes down to single cellular prokaryotes, and they all serve the same basic function – to assist nascent polypeptides in folding or to maintain them in an unfolded state. This can be achieved in a variety of ways depending on the particular chaperone and the ultimate location to which the substrate is destined. For example, the family of ribosome-associated chaperones, including but not limited to, trigger factor (TF) (*9-11*), heat-shock protein 70 (Hsp70) (*12*), and prefoldin (*13, 14*), act by stabilizing the ever-elongating nascent polypeptides directly at the site of synthesis, i.e., on the ribosome (*15*). Alternatively, chaperonin systems, such as DnaKJ/GrpE and GroEL/GroES, act in an ATP-dependent manner to directly assist newly synthesized preproteins to fold in the cytoplasm. While most nascent polypeptides (~65-80%) are thought to interact with trigger factor, a substantially smaller, but significant portion of proteins interact with DnaKJ (~10-20%) and even fewer with the GroEL/GroES (~10-15%) chaperonin system (*15*).

As might be expected, the ribosome-associated factors are the first chaperones to interact with nascent chains. For many small proteins that fold quickly on their own, this is the only assistance that they will receive during their lifetime. However, many larger proteins require additional assistance from other chaperonin systems, such as the ones mentioned above. While not all molecular chaperones have the same mechanism of action, in general, the heat-shock proteins, which are upregulated during periods of cellular stress, function to prevent protein aggregation by binding to exposed hydrophobic patches and unstructured backbone regions of their substrates – a hallmark of unfolded proteins (*15-18*).

Overall, there are many different molecular chaperones present within the cell of both prokaryotes and eukaryotes, all of which assist the cell in producing high quality, fully folded proteins in a well-orchestrated manner. These systems mark the first form of protein quality control and are a fundamental aspect of maintaining cell viability through the production of folded and functional protein products.

1.2 General protein secretion in bacteria

In bacteria, proteins are made by ribosomes in the cytoplasm and assisted by general molecular chaperones as described above, but not all proteins ultimately end up in the cytoplasm – some 20% are transported into the periplasmic space where they carry out their function (*19*). Bacteria have evolved elegant protein secretion pathways that are responsible for delivering protein substrates to their final destination. In general, there are two distinct pathways in *Escherichia coli*, the general secretory pathway (Sec) and the twin-arginine translocation pathway (Tat).

1.2.1 The general secretory (Sec) pathway

As the name suggests, the general secretory (Sec) pathway is responsible for the transportation of most secreted proteins across the cytoplasmic inner membrane in gram-negative bacteria, and across the cytoplasmic membrane and cell wall in gram-positive bacteria. Substrates targeted to this pathway are recognized by short (~20-30 amino acids in length) characteristic leader peptides known as Sec signal sequences. These sequences consist of three distinct domains: 1) a N-region that is largely basic in nature; 2) a hydrophobic domain of ~7-13 residues, termed the H-region; and 3) a slightly charged section known as the C-region (*20, 21*). With the exception of the above mentioned characteristics, these signal sequences are relatively similar to each other and lack many additional distinguishing features. Despite their generality, these short, cleavable (*22*) N-terminal extensions serve a very important function and are largely responsible for determining the final location of a protein.

In bacteria, nascent polypeptides synthesized with a N-terminal Sec signal sequence are detected by the ATPase motor protein, SecA (*23*). It was long thought that the Sec pathway was characterized by posttranslational translocation initiated by the pathway-specific molecular chaperone, SecB (*24*), while cotranslational translocation was achieved by the signal recognition particle (SRP) (*25*). However, recent studies have clearly demonstrated that SecA interacts with the ribosomal protein L23, at a location close to the polypeptide exit tunnel, suggesting that SecA could interact with its substrate *during* translation, i.e., in a cotranslational manner, in order to facilitate the efficient channeling of Sec substrates to this "posttranslational" pathway (*26*). Regardless of the specific molecular interactions of SecA, all current models of the general secretory pathway agree that Sec substrates are held in an unfolded confirmation and directed to the Sec translocase. This translocase consists of the three integral membrane proteins, SecY, SecE, and SecG, which form a complex

with a very narrow opening of only ~5-8Å in diameter that serves as the channel in which translocation can occur (*27*). The SecA molecule, a highly conserved and essential molecular motor, subsequently drives protein translocation through this very narrow pore in an unfolded confirmation through successive rounds of ATP hydrolysis (*28-31*). Only after the preprotein is present in the periplasmic space is it allowed to fold and achieve its final native confirmation.

Alternatively, while still utilizing the same SecYEG translocon, the signal recognition particle can also export proteins in an unfolded confirmation. This is achieved by immediately binding the nascent chain at the ribosomal exit tunnel (*32*) and directing the entire complex to the translocase, whereby the nascent chain is simultaneously exported as it is being translated, i.e., "cotranslational" translocation (*33*). It is important to note that the SRP pathway has a great deal of overlap with the Sec pathway, and even shares some of the same machinery. Furthermore, the common theme of this protein secretion pathway is that the substrate is maintained in an unfolded state, exported, and only allowed to fold after translocation has occurred.

1.2.2 The twin-arginine translocation (Tat) pathway

While the vast majority of secreted proteins in *Escherichia coli* are exported via the general secretory (Sec) pathway, approximately six percent (*34*) still exit the cytoplasm through a completely different system discovered approximately ten years ago. The hallmark of this fundamentally unique conduit, termed the **t**win-**a**rginine **t**ranslocation (Tat) pathway, named after its signature (S/T)**RR**xFLK signal peptide motif (*3, 35-37*), is its ability to allow only natively folded proteins passage across the cytoplasmic bacterial membrane, all the while maintaining delicate membrane integrity (*2, 38-41*). This is in stark contrast to both the SRP and Sec pathways, which only transport unfolded substrates.

The Tat pathway was first discovered in plant chloroplast thylakoid membranes as a ΔpH-dependent protein import system (*42, 43*). Since then, analogous protein export pathways have been identified in both bacteria and archaea. Interestingly, the Tat pathway also has been recently implicated in the pathogenicity of many bacterial pathogens, e.g., *Agrobacterium tumefaciens* (*44*), *Escherichia coli* O:157 (*45, 46*), *Legionella pneumophila* (*47, 48*), and both *Pseudomonas aeruginosa* (*49*) and *syringae* (*50*). As a likely result of the potentially horrifying consequences to human health, there have been great strides in the past several years to better understand the nuances of this pathway.

In gram-negative bacteria, the Tat machinery consists of three membrane-bound proteins: TatA (paralogue TatE), TatB, and TatC. While TatA, TatB, and TatE form single transmembrane α-helices, followed by another short, cytoplasmic helix at their C-termini, TatC is structurally much more complex, consisting of six transmembrane α-helices (*51-56*). Furthermore, deletion of any of these proteins will partially (in the case of TatA or TatE alone) or completely (in the case of TatB, TatC, or TatA and TatE together) block export of Tat-dependent substrates (*37, 57-59*), giving rise to a variety of characteristic phenotypes, e.g., formation of long filamentous chains and membrane defects (*60*).

The current working model of Tat translocation suggests that the substrate protein may be directly inserted into the lipid membrane as the initial step (*61, 62*), followed by an interaction with a receptor complex, consisting of the essential proteins, TatB and TatC (*57, 58*). Regardless of how membrane insertion occurs, the TatBC pre-complex then associates with the homo-oligomeric TatA complex, forming a transient TatABC translocase, which is responsible for exporting the natively folded substrates into the periplasm (*63, 64*). However, the precise mechanism in which it is able to perform this feat is unknown. One possibility is that the TatA homo-

oligomeric complex forms a ring structure around the substrate, allowing transport of the folded substrate to occur through its pore. While several single-particle imaging studies have documented the formation of a double ring-like structure with Tat-specific substrates bound (*65, 66*), these studies fail to adequately describe how the Tat translocase can accommodate such a vast array of diameters due to the export of native substrates ranging from ~10-140kDa in size (*3*). Alternatively, a second possibility was proposed whereby the TatA homo-oligomeric complex would act to locally destabilize the membrane, allowing for substrates to be subsequently pulled through it (*67, 68*). Regardless of which method, if either is correct, in contrast to the ATP-dependent Sec pathway, the export of Tat substrates is known to occur by the proton motive force characteristic of charged membranes (*69*), whereby the energy needed to drive export is acquired solely by the exchange of protons. Once the substrate is in the periplasm, the Tat signal sequence is cleaved by signal peptidase I and the mature protein is released, marking the end of its journey.

Due to the distinct virulence role of the Tat pathway (*70*), elucidating the molecular mechanisms driving this system have been a significant focus of many research studies with the goal of developing potential downstream biotechnological applications, e.g., broad-spectrum antibiotics. However, much more interesting is the idea of actually harnessing the pathway's exceptional ability to transport only folded proteins, and subsequently, engineering novel protein architectures that have passed through this cellular pre-screen. Several examples of this have already come to fruition, such as, the engineering of robust single-chain variable fragment (scFvs) antibodies (*71*) and enhancing the solubility of aggregation-prone substrates (*72*). Nevertheless, understanding the intrinsic mechanisms that drive this quality control will truly be the key to unlocking the most rewarding downstream engineering applications.

1.2.3 Quality control along the Tat pathway in **Escherichia coli**

The mere notion that folded proteins are exported via the Tat machinery, while unfolded or otherwise misfolded substrates are not, suggests the presence of some form of quality control (QC). However, there currently exists very little evidence about the molecular mechanism by which the Tat system is able to perform this most remarkable task. According to the above mentioned model of Tat export, it is conceivable that quality control could occur at any of these stages.

While some studies have shown accessory factors involved in substrate proofreading in a chaperonin-like manner (*73-76*), these studies are substantially less interesting due to obvious similarities with the Sec pathway and general molecular chaperones. In particular, we previously discussed a general class of broad-acting molecular chaperones with very little substrate specificity. In addition to these promiscuous chaperones, there also exists many substrate-specific chaperones unique to particular protein translocation pathways, and even, to individual protein substrates. The more general of these chaperones are typically found in the general secretory pathway, such as SecA and SecB. Conversely, the twin-arginine translocation pathway has a number of cytoplasmic chaperones, such as, DmsD and TorD (*74, 77*), which bind only a few unique substrates. While these chaperones are largely substrate specific, they function in much the same way as general chaperones in that they bind to the unique Tat signal sequence of their cognate substrates, sequester them away from the Sec machinery, assist them in folding, and ultimately direct them to the Tat translocon (*78*).

Nevertheless, a growing body of biochemical studies (*79, 80*), suggest that the TatB and TatC proteins form a receptor complex early in the translocation process, with which both folded and misfolded Tat substrates interact and potentially facilitate Tat-mediated folding quality control. Furthermore, additional biochemical studies

have shown that the TatBC complex, specifically the N-terminus of TatC, acts to initially recognize the signature Tat signal sequence (*81-83*), suggesting that the Tat machinery itself indeed has a pivotal quality control responsibility early on in the translocation process. Moreover, it has been recently shown *in vitro* that the Tat-competent substrate, oxidized ssTorA-PhoA, cross-linked to the TatBC complex (*81*) in a completely different manner as compared to its Tat-incompetent form (i.e., reduced), suggesting that the Tat machinery was capable of discerning the folding state of the substrate (*80*). Interestingly, it also has been recently shown that the Tat apparatus initiates degradation of misassembled FeS protein substrates, as evidenced by NrfC/NapG degradation in wild type strains but accumulation in either *tat* null strains or Tat-misdirected strains (*84*), providing further indication of its innate proofreading capabilities.

Taken together, the fact that both transport-competent and transport-incompetent substrates are targeted to the Tat translocase, yet only passage of properly folded/assembled substrates is allowed, strongly favors the hypothesis that Tat-mediated folding quality control is a function of the machinery components themselves. Therefore, the true merit of these studies lies in the fact that no other system in biology uses the export machinery itself in proofreading and quality control, and thus it would distinguish the Tat machinery as a unique example in all of biology. Hence, the research conducted herein will focus on the elucidation of the molecular mechanisms responsible for this putative inbuilt quality control feature.

1.3 Protein degradation pathways in Escherichia coli

An inevitable byproduct of protein folding, is protein misfolding. As such, organisms have necessarily evolved machinery components capable of breaking down malformed, dysfunctional proteins into their constituent elements, i.e., amino acids,

for reuse. These degradation systems are the final form of quality control within the cell, whereby misfolded, aggregated, or otherwise unusable protein is rapidly degraded by precise molecular machines, cleared away to reduce clutter within the cell, and ultimately recycled.

In bacteria, several examples of these degradation systems include, but are not limited to, FtsH, Lon and the Clp degradation machinery (*85, 86*). In the context of exported proteins, it is also noteworthy to mention the periplasmic protease, DegP, which is likewise responsible for degrading misfolded periplasmic proteins (*87, 88*). Interestingly, proteases largely act via a similar mechanism as molecular chaperones, in that, they recognize solvent exposed hydrophobicity in misfolded substrates, bind them, and subsequently perform their natural function – in this case proteolysis (*89*). Therefore, while proteases function as a general housekeeping mechanism within the cell, they are still extremely important in maintaining overall quality control along the protein folding landscape.

In summary, we have discussed a variety of protein folding quality control mechanisms present within the cellular environment. These systems range from accessory factors participating early on in the translation process and assisting newly synthesized proteins to fold, to protein secretion pathways, and ultimately to degradation pathways invoked upon protein misfolding. The most interesting of these processes, i.e., Tat-mediated folding quality control, will be discussed further in the remaining chapters.

CHAPTER 2

EXPLOITING THE INBUILT FOLDING QUALITY CONTROL MECHANISM OF THE TAT MACHINERY PROTEINS TO DISCRIMINATE BETWEEN FOLDED AND MISFOLDED SUBSTRATES

2.0 Introduction

While a variety of studies have demonstrated that the Tat system can discriminate between folded and misfolded substrates (*38, 71, 72, 90*), there remains to be a study that has developed such a system capable of decoupling the terms "unfolded" and "unstable/insoluble," due to the inclination of the majority of misfolded proteins to form insoluble precipitates and/or become rapidly targeted for degradation. Thus, the pursuit for the ideal target substrate that will allow for a thorough dissection of the Tat pathway's proofreading and quality control mechanisms remains a coveted prize. Recently, we have found such a target with the *de novo* designed three helical bundle proteins, termed α_3A, α_3B, α_3C, and α_3D (*91, 92*).

The benefit of using this family of proteins as a model substrate-set lies in the very nature of *de novo* design. With significant advances in the field of bioinformatics and the ever expanding power of computational analyses, the world of computer-based algorithms has given birth to the field of *de novo* protein design (*93, 94*). *De novo* design embodies the idea of customizing a functional protein of distinct structure from scratch. Several recent studies have successfully created generic, quantitative computational algorithms based on protein structure determining physicochemical properties (*95*), enabling the *de novo* rational design of simple proteins with well-ordered secondary and tertiary structures (*91, 93, 96, 97*). Moreover, a myriad of structurally important features have been computationally predicted, and subsequently

created, including functional metal-binding pockets (*98, 99*), zinc fingers (*100-102*), coiled coils (*103*), and a variety of additional small protein domains (*104-107*). As such, it becomes possible to create and utilize native-like small proteins of minimally sufficient secondary and tertiary structure catered to a specific need.

To this end, we selected the highly thermostable α_3D protein because it perfectly fills our need for a well-characterized and well-ordered model protein of clear directed evolutionary lineage. This thermostable three helical bundle originates from a similarly *de novo* designed antiparallel three-stranded coiled-coil, termed Coil-Ser (*108*). The α_3D variant is the culmination of multiple, iterative rounds of design, each resulting in a more evolved species. In particular, the final ordered, thermostable structure was achieved through the addition of helical capping boxes (*109, 110*), rearrangement of select residues allowing for stabilizing electrostatic interactions to occur (*111*), and the repacking of the hydrophobic core with a diversified set of amino acids (*112*). α_3D was shown to display thermodynamic and spectroscopic properties typical of native proteins, validating the *de novo* prediction and design. The precise modifications at each step that produced each successive α_3 protein are enumerated in Table 2.1 below as determined by (*91, 92*).

Table 2.1 Features of the *de novo* designed proteins α_3A-D.

Variant	Structure	Features	ΔG_D (kcal/mol)
α_3A	Molten Globular	Parental platform: Coil-Ser (Ref. 108) Hairpin loops added between helices - failed to stabilize bundles Aggregation prone - formation of monomer/dimer/trimer structures	---
α_3B	Molten Globular	Addition of helical capping boxes Able to form a defined 3 helical bundle Possessed molten globular characteristics Lacked a nonpolar core - essential for functional proteins Minimized aggregation - predominantly monomers	7.2
α_3C	Ordered	Repacked and rearranged the hydrophobic core Near "native-like" - ΔC_p, GdnHCl denaturation, ANS binding Monomeric	5.5
α_3D	Ordered	Diversification of amino acid identity (19 of 20 aa used) Native like character compared to natural proteins Monomeric	5.1

These three helical bundle proteins are aligned below to show the similarities in their primary sequences and the rationally designed modifications made between each iterative round of design (Figure 2.1a and Table 2.1). Moreover, the NMR solution structure of the most thermodynamically stable variant, α_3D, was solved (*92*) to confirm the desired tertiary structure that was designed *de novo* (Figure 2b&c).

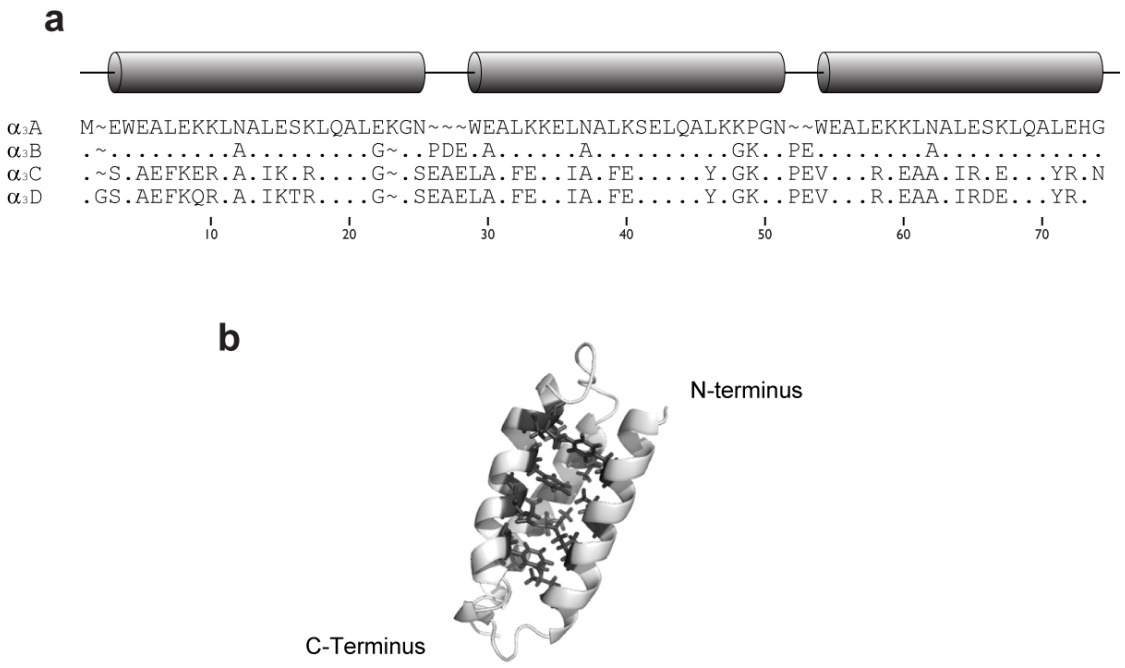

Figure 2.1 Primary and tertiary structures of the α_3 proteins. (a) Primary sequence alignment of the *de novo* designed three helical bundle proteins, α_3A, α_3B, α_3C, and α_3D – all rationally designed from the same parental Coil Ser *de novo* protein. Predicted helices are drawn above the primary sequences at the appropriate locations. (b) Three-dimensional solution structure of α_3D (PBD ID: 2a3d) as determined by Walsh *et al.* (1999) (*92*), displaying the repacked core of hydrophobic sidechains shaded dark gray.

Taken together, the highly thermostable α_3D and its less evolved precursor α_3's of the same family provide a perfect continuum of folded structure, ranging from molten globules (α_3A and α_3B) to well-ordered three helical bundles (α_3C and α_3D). Consequently, herein we report the use of this well-characterized family of proteins to

genetically dissect the quality control feature inherent in the Tat pathway with much greater resolution than was previously possible.

2.1 α_3 de novo *proteins as a Tat-mediated screen for folding*

As a fortuitous side effect of the quest for a rationally designed, well-ordered, three-helical bundle, with native-like folding character, DeGrado and coworkers created a family of *de novo* proteins with progressively increasing order and structure (*91, 92*). Furthermore, despite varying degrees of folding, all four variants remained cytoplasmically soluble and could be detected by immunoblotting. Thus, these *de novo* designed proteins, termed α_3A through α_3D, became the perfect tool for interrogating the putative quality control mechanism of the twin-arginine translocation pathway. Since all α_3 variants are soluble at comparable levels, we were able to exploit their differential thermal stabilities in an effort to develop a functional screen of substrate folding character, effectively decoupling "soluble" from "folded" for the first time. This feat was accomplished by expanding on the pSALect-based selection for protein solubility described by Fisher *et al.* as a tool to correlate the cell growth on selective media with the solubility of a target protein (*72*).

In particular, the DNA sequence encoding each α_3 protein substrate was optimized for *E. coli* expression and cloned into the pSALect vector – in between a N-terminal TorA signal sequence (ssTorA) and a C-terminal TEM-1 β-lactamase reporter – to create a tripartite fusion protein (*113*). Under the control of the *lac*-promoter, the translated protein fusion is directed to the Tat machinery due to the TorA signal sequence, and if folded, exported across the inner membrane and into the periplasm where the C-terminal TEM-1 β-lactamase fusion partner can hydrolyze β-lactam ring derived antibiotics, such as ampicillin or carbenicillin, conferring resistance to said antibiotics. Therefore, with the solubility levels of each α_3 variant nominally uniform,

using ampicillin as a selective marker, protein substrates may be screened for folding character alone as indicated by cellular resistance to ampicillin – the greater the resistance, the better folded the substrate protein.

In an effort to develop this system into a functional screen for substrate folding character, constructs pSALect–α_3A, B, C, and D in a *wild-type* MC4100 strain background were serial diluted and spotted onto LB-agar plates supplemented with 200μg/ml of ampicillin and allowed to grow overnight at 30^{o}C. Strikingly, cells harboring the α_3D substrate grew out to a dilution of 10^{-6}, whereas, identical cells carrying the α_3A variant only grew to a dilution of 10^{-1}. In agreement with their respective *de novo* structures, cells with the α_3B and α_3C variants grew to intermediate dilution factors (Figure 2.2a, left panels), suggesting that rationally designed proteins of varying order and stability can be differentially exported via the Tat pathway.

However, with these findings alone, it is still uncertain as to whether the Tat machinery itself is responsible for this quality control. If it is so, this would be the first of its kind in biology and an incredible avenue for engineering recombinant proteins produced through the Tat pathway. To begin probing the answer to this question, *tatC* knockout cells (*58*) were transformed with the same pSALect–ssTorA–α_3–TEM1 fusions and spotted on ampicillin containing LB-agar plates. TatC is an essential component of Tat-mediated translocation, and thus, *tatC* knockouts will completely abolish transport through this pathway (*58*). Therefore, if this growth phenotype is truly a product of Tat quality control, these Δ*tatC* cells will not transport the β-lactamase fusions and will not grow on selective media. Indeed, the growth phenotype was abolished in Δ*tatC* cells harboring any variant of the α_3–β-lactamase fusions (Figure 2.2a, right panels), suggesting that this observation is in fact a result of a tat-specific phenomenon.

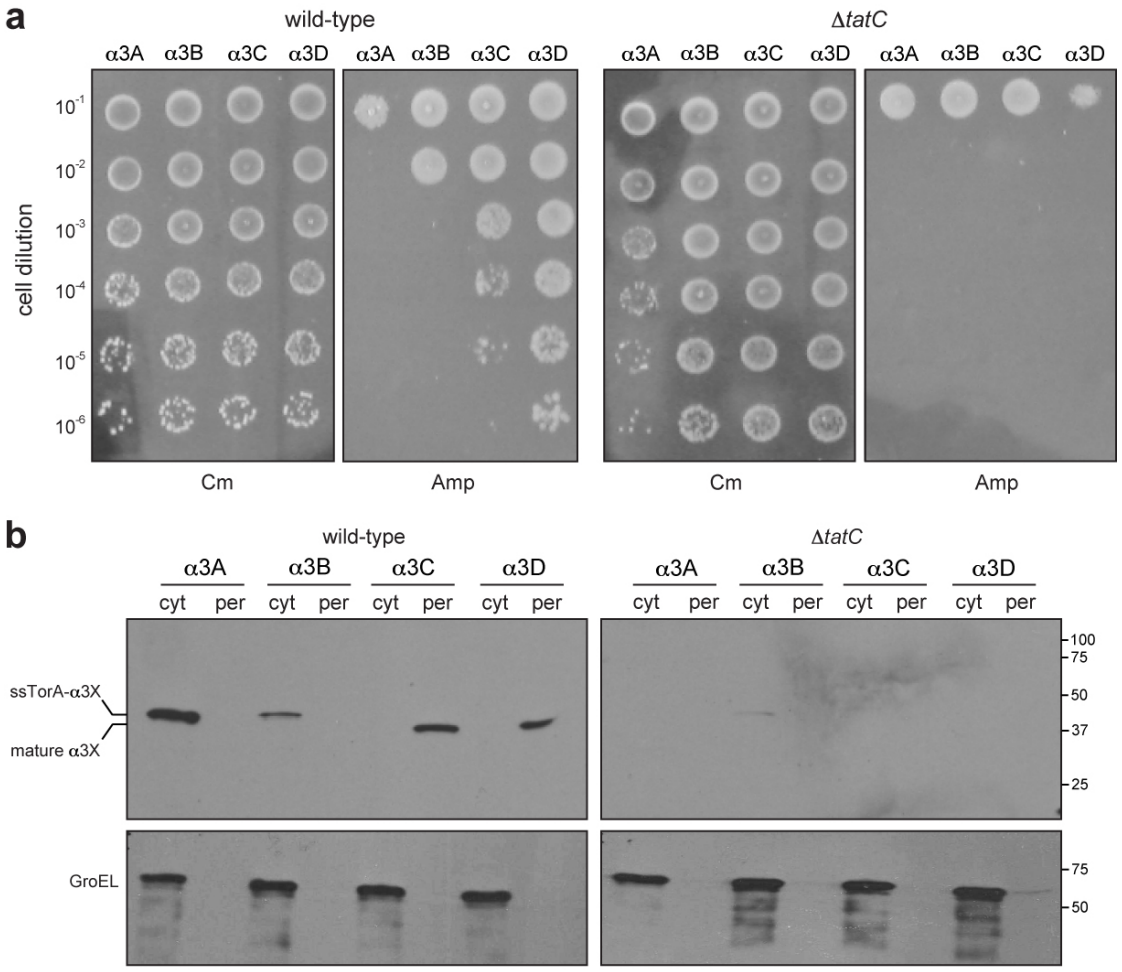

Figure 2.2 Genetic demonstration of the inbuilt folding quality control mechanism of the Tat machinery proteins. Establishment of a robust genetic screen for Tat-mediated quality control. (a) As indicated, cells containing each protein-fusion were serial diluted and spotted onto plates containing either 25 μg/ml chloramphenicol or 200μg/mL ampicillin, respectively. Wild-type MC4100 cells, harboring an intact Tat pathway differentially export the protein-fusion, consistent with the folded state of the target α_3 protein. Conversely, B1LK0 cells (Δ*tatC*), deficient in Tat export, show a marked deficiency in growth, demonstrating Tat-mediated exclusivity. (b) Western blot analysis of cytoplasmic and periplasmic fractions prepared from wild-type MC4100 cells (left) and Δ*tatC* B1LK0 cells (right) by the ice-cold osmotic shock protocol and blotted against α-Bla (mouse monoclonal 1ºα-Bla diluted 1:100 and goat α-mouse-HRP diluted 1:2,500), demonstrating the localization of the protein-fusion in the cytoplasm for the molten globular target α_3s and in the periplasm for their ordered counterparts. Consistent with a Tat-specific housecleaning mechanism, all protein-fusions in the Δ*tatC* strain are almost entirely degraded due to the genomic deficiency in Tat export.

In conjunction with the β-lactamase spot plate screen, subcellular fractions were prepared via the ice-cold osmotic shock procedure (*38*), followed by Western blot analysis in order to confirm the presence of soluble substrate in the periplasm of transport-competent cells. Trimethylamine N-oxide reductase (TorA) is an exclusive Tat substrate (*38*) limiting the export of β-lactamase to the Tat pathway in the form of a fusion partner to the α_3 protein substrates. Therefore, β-lactamase will only be present in the periplasm if its fusion partner, i.e. α_3A, B, C, or D, is properly folded and competent for transport. As expected, the presence of β-lactamase was detected in the periplasm for the α_3C and α_3D substrates and in the cytoplasm for the molten globular cases of α_3A and α_3B (Figure 2.2b), corroborating the *de novo* design work of DeGrado *et al.* (*91, 92*) and the protein solubility selection of Fisher *et al.* (*72*). Furthermore, in agreement with the current knowledge-base of Tat research, there was little to no detectable protein in either the cytoplasm or the periplasm in a Δ*tatC* strain background. It has been observed by a variety of scholars that an active "housecleaning" mechanism associated with the Tat pathway rapidly degrades transport-incompetent substrates (*38, 72, 114*), and thus, it is no surprise that all variants are no longer detectable in this strain background.

2.2 Discussion

In this study we have developed a genetic means of decoupling unfolded substrates from insoluble substrates, by utilizing a family of *de novo* designed proteins of increasing order, but comparable soluble expression. In doing so, we have also demonstrated that it is the Tat machinery proteins themselves responsible for discriminating between folded and misfolded substrates. As mentioned earlier, the true scientific merit of this study lies in the fact that no other system in biology enrolls the participation of the export machinery itself in proofreading and quality control, and

thus, the evidence presented in this study acts as the first direct experimental evidence supporting these claims. Furthermore, from an engineering perspective, the research described in this chapter has focused on first understanding this unique and complex system on a basic science level in order to eventually exploit the intrinsic folding quality control feature of the Tat pathway, as will be described in the remaining chapters.

Due to its potential to ensure the quality and structural integrity of exported proteins, elucidating the molecular mechanisms driving this system hold significant promise for the development of downstream biotechnological applications, e.g., engineering recombinant antibodies. However, a major obstacle in the development process has been the lack of quality genetic screens aimed at understanding this Tat-based quality control. Thus, it is of paramount importance to develop an elegant genetic selection capable of screening recombinant libraries. With advances in error-prone PCR, a library selection is a convenient, high throughput method to explore the functional diversity of a vast sequence space of a desired gene or gene cluster, thereby exhausting an immense number of possible amino acid sequences responsible for specific gene functions. As such, the goal of the work presented in this chapter was to lay the foundation for the eventual development of a robust genetic selection of Tat-mediated quality control that will ultimately be capable of screening large recombinant libraries of the *tatABC* operon in order to isolate broad-spectrum mutants suppressing Tat quality control. The details of such a genetic selection will be presented in the following chapter.

2.3 Materials and methods

2.3.1 Bacterial strains, plasmids and growth conditions

MC4100 (F^- Δ*lacU169 araD139 rpsL150 relA1 ptsF rbs flbB5301*) (*115*) served as the parental *E. coli* strain for all experiments conducted in this study. An isogenic *tatC* knockout strain, B1LK0, was also used as indicated. All genes encoding the α_3 family of proteins were codon optimized for expression in *E. coli* and cloned into the protein-solubility selection plasmid, pSALect, using the NdeI and SpeI restriction enzyme sites as described (*72, 113, 116*). Below is a full list of all strains and plasmids used in this study. All plasmids generated in this study were confirmed by sequencing.

Table 2.2 Strains and plasmids used in this study.

Strain or Plasmid	Description	Reference
MC4100	*F- ΔlacU169 araD139 rpsL150 relA1 ptsF rbs flbB5301*	Ref. 115
B1LK0	MC4100 derivative lacking the *tatC* gene	Ref. 58
pSALect	Universal inframe cloning vector encoding a TorA signal sequence fused to TEM-1	Ref. 113
pSALect-α3A	The gene a3A cloned into the POI position using the NdeI and SpeI cutsites	This work
pSALect-α3B	The gene a3B cloned into the POI position using the NdeI and SpeI cutsites	This work
pSALect-α3C	The gene a3C cloned into the POI position using the NdeI and SpeI cutsites	This work
pSALect-α3D	The gene a3D cloned into the POI position using the NdeI and SpeI cutsites	This work

2.3.2 Spot plating and growth studies

Spot plating was performed as previously described (*72, 116*). Briefly, MC4100 or B1LK0 cells harboring a pSALect–ssTorA–α_3–TEM1 plasmid were grown overnight at 37ºC in Luria Bertani (LB) medium supplemented with 25 μg/ml

chloramphenicol. One hundred microliters of cells were removed from each culture, resuspended in fresh medium, and subsequently serial diluted by factors of ten in a 96-well tissue culture plate. Five microliters of each sample were spotted onto ampicillin containing LB-agar plates until all samples at all dilutions were accounted. Spot plates were incubated overnight at 30ºC for approximately 16.5 hours.

2.3.3 Subcellular fractionation and Western blot analysis

Overnight cultures were prepared as described above. The following day, overnight cells were subcultured ten-fold in LB containing antibiotics and allowed to grow for an additional 90 min in fresh medium at 37ºC until an optical density of OD_{600} ~0.5 was reached, at which time the cultures were induced with 1.0 mM IPTG and transferred to a 30ºC incubator. For pSALect-based substrates, protein expression proceeded at 30ºC overnight for approximately 16.5 hours, at which time the cultures were pelleted via centrifugation for 15 min at 4ºC and 3,500 rpm. Cell pellets were subsequently fractionated according to the ice-cold osmotic shock method as described elsewhere (*38*). Induction conditions were performed in this manner to remain consistent with spot plate growth experiments. Proteins were separated on a 12% polyacrylamide gel (Bio-Rad) by SDS-PAGE and detected by Western blotting also as previously described (*117, 118*). Protein detection was determined with the following primary antibodies: mouse monoclonal 1ºα-Bla (Abcam) diluted 1:100 and rabbit polyclonal 1ºα-GroEL (Abcam) diluted 1:10,000. The following secondary antibodies were used as appropriate: goat α-mouse-HRP (Promega) diluted 1:2,500 and goat α-rabbit-HRP (Promega) diluted 1:2,500.

CHAPTER 3

DEVELOPMENT OF A GENETIC SELECTION FOR THE ISOLATION OF MUTANT TRANSLOCASES SUPPRESSING QUALITY CONTROL

3.0 Introduction

The unique ability of the Tat pathway to transport only *folded* substrates across the bacterial inner membrane has sparked a vicious debate as to precisely how it is able to discern folded from misfolded. There are generally two distinct possibilities that have arisen in response to early observations (*38*) suggesting that the Tat pathway processes both folded and misfolded substrates, but transports only the former. The first possibility is the existence of one, or many, accessory factor(s) that perform a chaperonin-like function, binding to and sequestering substrates away from the machinery until they have achieved a threshold fitness to be exported (*73, 74, 79, 119*). Alternatively, a much more interesting possibility is that the Tat machinery proteins, i.e., TatA, TatB, and TatC, themselves are responsible for such scrutiny, granting transport approval for those deemed fit enough to pass quality control and rejecting those that are unfit for passage (*38*). While first hypothesized in the seminal work by DeLisa and coworkers, there remains to be any direct, incontrovertible genetic evidence in live bacterial cells suggesting that this latter possibility is in fact the case. Despite the lack of genetic evidence, there are several biochemical studies demonstrating that both folded and misfolded substrates do seem to interact with the Tat proteins, and thus, suggest that they could indeed have this ability.

As such, it is of great interest to help elucidate the initial quality control steps along the Tat pathway, and especially, to determine the extent to which the Tat machinery itself undertakes this role. Furthermore, given the nature of the transport

phenomenon, which minimally involves three distinct components that all seem to interact together (i.e. TatB, TatC, and the substrate), it is particularly prudent to undertake this feat in an unbiased manner, akin to library selection. Specifically, it is entirely possible that conformational changes of both TatB and TatC are required for binding, recognition, and/or transport of a substrate through the Tat pathway. Moreover, recent evidence suggests that not only the TatABC complex (*82*), but also merely the TatBC complex undergoes dramatic conformational changes upon substrate binding (*120, 121*), further bolstering the need for a *tatABC* library as opposed to simply a *tatBC or tatC* library. Accordingly, screening a randomly mutated library of the entire *tatABC* gene cluster would provide the perfect avenue to investigate suppressor mutations, allowing for the isolation of clones with compensatory mutations, that is, two or more mutations that interact with each other, restoring a potentially adverse effect of only a single mutation (*122*). This is in part the rationale for co-evolution studies aimed at evolving a particular enzyme in the context of how it interacts in its native environment.

Taken together, a genetic study linking changes in the actual genetic code to differential Tat-mediated export would serve at the ideal evidence needed to support the quality control hypothesis. To this end, we sought to create a randomized library of mutant *tat* translocons and select for clones that would relax this putative inbuilt quality control feature, and doing so, would provide the first concrete genetic evidence directly linking the Tat machinery proteins to folding quality control.

3.1 Isolation of a broad class of mutant TatABC translocases capable of suppressing quality control

As a preliminary, yet crucial, step toward the development of a *tatABC* library was the verification of the characteristic growth phenotype and selective periplasmic export of α_3 substrate chimeras from a *tatABC* plasmid expressed construct. To this end, pTatABC was constructed from a pBR322 vector. In particular, the entire *tatABC* operon, which encodes the single polycistronic mRNA strand giving rise to the TatABC proteins (*37*), was cloned into the pBR322 backbone in place of the ampicillin resistance marker, resulting in only a tetracycline resistance cassette. Next, DADE (MC4100 Δ*tatABCD* Δ*tatE*) cells (*123*) were co-transformed with both the pTatABC (harboring the *tatABC* gene cluster) construct and each pSALect–ssTorA–α_3–TEM1 partner plasmid, followed by spot plating on ampicillin selective media and Western blot analysis of its subcellular components. Indeed, plasmid expressed TatABC proteins produced the same characteristic phenotype (Figure 3.1a, leftmost panel) as genomic TatABC, albeit at a greater ampicillin concentration. However, this is not uncommon due to the differences in copy number and expression between genomic- and plasmid-based sources. Furthermore, the identical transport phenomenon was observed via Western blotting (Figure 3.1b leftmost panel), confirming the functionality of the plasmid encoded TatABC machinery and providing an adequate template for randomized mutagenesis.

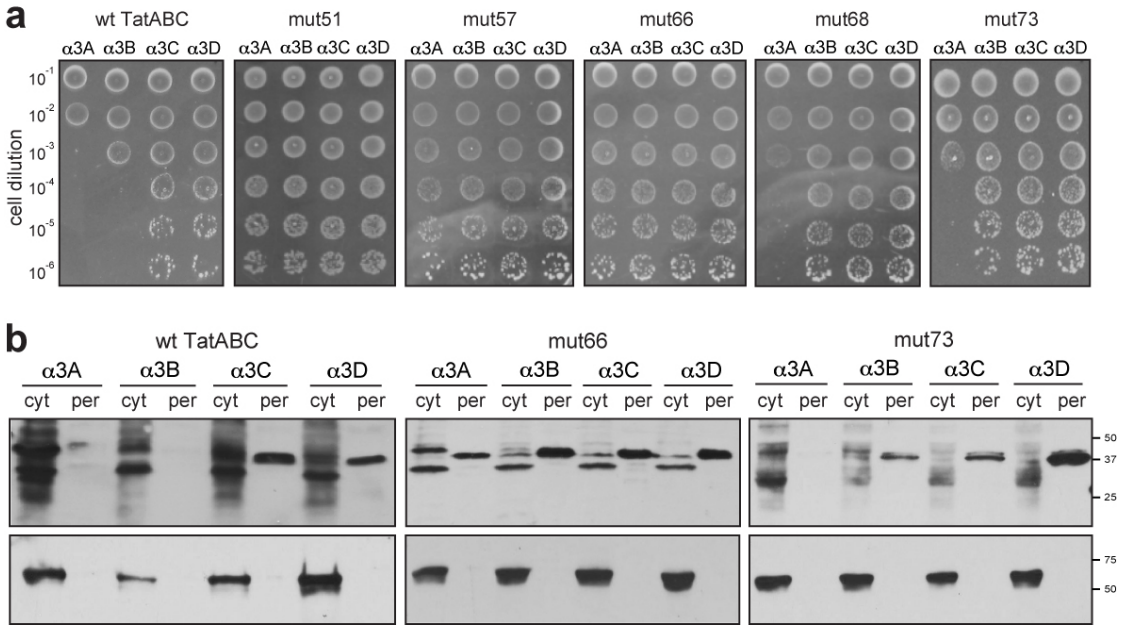

Figure 3.1 Isolation of class I and class II suppressor translocases based on growth on ampicillin and Western blot export of the α_3 substrates. A counter-screen of mutant translocases reveals suppressors of either the α_3B–TEM1 fusion alone or both α_3A– and α_3B–TEM1 fusions. The top five suppressor mutants were transformed into DADE (Δ*tatABCDE*) cells harboring each of the pSALect–ssTorA–α_3–TEM1 plasmids, respectively, serial diluted by factors of ten, spotted onto 500μg/mL ampicillin containing plates, and incubated overnight at 30ºC. Class I suppressors, comprised of mutants 51, 57 and 66, were capable of transporting the least ordered, molten globular α_3A protein-fusion. Conversely, class II suppressors, namely mut 68 and mut 73, were only capable of exporting the selected molten globular substrate fusion, α_3B–TEM1. Analysis of α_3 export was based on both growth on selective media (a) and Western blot analysis of subcellular fractions (b). Western blots characteristic of each suppressor class are shown and were performed as described in the Materials and methods section.

Error-prone PCR (*124*) was used to generate a random *tatABC* library of ~5.5 x 10^6 members with an average random error rate of 2.0 nucleotides per gene. The library was transformed into DADE (Δ*tatABCDE*) cells already harboring the pSALect–ssTorA–α_3B–TEM1 plasmid and selected on an ampicillin concentration of 200 μg/ml. The α_3B molten globular substrate was chosen due to its lack of export, localization in the cytoplasm, and ability to remain soluble without an inclination to aggregate (*91*).

Thirty randomly chosen clones were isolated, cured of their pSALect plasmid, back-transformed into fresh DADE (Δ*tatABCDE*) cells containing the pSALect–ssTorA–α_3B–TEM1 plasmid, counter-spot-screened on 500μg/ml of ampicillin and rank-ordered according to their ability to grow on ampicillin selective media. There were a variety of growth patterns for all of these mutants with more than ten growing out to a dilution factor of 10^{-6}, similar to that of the well-ordered, native-like three-helical bundle, α_3D. The top five growers were further characterized to determine the extent of suppression. In particular, each pTatABC mutant was freshly co-transformed into DADE (Δ*tatABCDE*) with each α_3–TEM1 fusion encoded on the pSALect plasmid and spotted onto ampicillin selective plates (Figure 3.1a). Strikingly, three of these five mutants, mut 51, mut 57, and mut 66 (hereafter referred to as class I suppressors) were even capable of transporting the aggregation prone, molten globular α_3A–TEM1 fusion (Figure 3.1a, leftmost panels), suggesting that these are indeed genuine suppressors of quality control and that the Tat machinery is in fact performing this suppression such that translocation of previously transport-incompetent substrates can occur. To our knowledge, this is the first direct genetic evidence linking mutations in the Tat machinery to altered export of substrate proteins based solely on their folding character. Equally interesting is the fact that all three of these class I suppressors have very different mutations in all three genes, confirming the hypothesis of necessary compensatory mutations enabling functionality (*122*). Conversely, mut 68 and mut 73, comprising a second class of best putative suppressors (hereafter referred to as class II suppressors), were fully capable of exporting the selected fusion substrate, α_3B–TEM1, as determined by growth out to a dilution factor of 10^{-6} on 500μg/ml of ampicillin; however, they were unable to tolerate the less-ordered molten globular fusion, α_3A–TEM1 (Figure 3.1a, rightmost two panels).

Although we have previously shown in an MC4100 parental strain background that growth on ampicillin plates directly correlated with subcellular localization of the target substrate, the first question we asked of the top five growers post-selection was whether the substrate fusion could also be detected in the appropriate subcellular compartments. Indeed, all mature α_3–TEM1 fusions were found localized in the periplasm for the class I suppressors, while only α_3B, C, and D were exported into the periplasm for the class II suppressors (Figure 3.1b). From a wild-type translocase to class I and class II suppressors, the growth phenotype seen on ampicillin containing plates directly correlated with the Western blot localization of the substrate.

3.2 Discussion

In this study we have utilized the genetic means of decoupling "unfolded" from "insoluble" substrates developed in Chapter 2 to screen a randomly mutated *tatABC* library containing ~5.5 x 10^6 members and directly select mutants allowing growth at a concentration that previously caused total cellular death. The ability to grow on higher ampicillin concentrations alone does not necessarily indicate that these clones have suppressed quality control. Rather, it merely demonstrates that the mutations introduced into the machinery components have allowed for more TEM-1 fusion to reside in the periplasmic space. This could occur one of three ways. First, the mutations could result in a hyperactive translocase that grows at higher ampicillin concentrations simply because it is able to export the substrate at a faster rate. Second, the mutants could have somehow destabilized the membrane integrity, allowing every substrate passage across the inner membrane. Third, the mutations could result in genuine suppressors of quality control. That is, our original hypothesis would be validated. However, if either scenario 1 or scenario 2 were the case, it would be unlikely that the substrates would be exported at levels comparable to a wild-type

translocase as determined by the Western blot analysis conducted within this chapter. Instead, one or both of the following two observations would be expected, but were not observed: 1) the mutants would export a much larger quantity of substrate protein into the periplasm due to their new hypersecreting ability; and 2) cytoplasmic chaperones, such as GroEL or DnaKJ, would leak into the periplasmic fraction due to the aforementioned membrane defects.

Given the evidence presented in this chapter, the third alternative is the most promising option and will be extensively explored in the following chapter as there are a number of questions still remaining that would validate this result. For example, several interesting questions that would aid in the classification of these mutants and overall confirmation of authentic quality control suppression would be as follows: How do the mutants behave in the context of native Tat substrates? How do the mutants respond to other non-native Tat substrates, those simply re-directed to the Tat pathway, or those presumably misfolded due to re-routing? Are there growth defects as compared to wild-type Tat machinery strains? Do the mutants retain some form of QC, or has all semblance of quality control been abolished? In other words, has quality control been truly suppressed, or completely eliminated?

Taken together, this study provides the first direct genetic evidence suggesting that the Tat machinery components themselves indeed possess an inbuilt folding quality control mechanism. While the precise molecular nuances responsible for this suppression still remain elusive, the development of a quality genetic selection capable of isolating mutants with relaxed folding quality control holds great promise for the future elucidation of these molecular details.

3.3 Materials and methods

3.3.1 Bacterial strains, plasmids, and growth conditions

MC4100's isogenic *tat* null strain, DADE (Δ*tatABCDE*) served as the parental *E. coli* strain for all experiments conducted in this study. The parental plasmid-encoded Tat machinery template was created by PCR amplification of the entire *tatABC* operon from MC4100 and cloned into the pBR322 backbone in place of the ampicillin resistance marker, resulting in a vector with only a tetracycline resistance cassette. More specifically, the *tatABC* operon and its native promoter were ligated into pBR322 using the restriction sites AhdI and PvuI, resulting in the pTatABC vector. This vector was subsequently used to create all downstream mutant variations used in the genetic selection described below in detail. Below is a full list of all strains and plasmids used in this study. All plasmids generated in this study were confirmed by sequencing.

Table 3.1 Strains and plasmids used in this study.

Strain or Plasmid	Description	Reference
DADE	MC4100 derivative lacking all *tat* genes	Ref. 123
pTatABC	pBR322 backbone encoding the wild-type *tatABC* genes with its native promoter; Tc^R	This work
pTatABC-Mut51	As pTatABC with quality control suppressing mutations	This work
pTatABC-Mut57	As pTatABC with quality control suppressing mutations	This work
pTatABC-Mut66	As pTatABC with quality control suppressing mutations	This work
pTatABC-Mut68	As pTatABC with quality control suppressing mutations	This work
pTatABC-Mut73	As pTatABC with quality control suppressing mutations	This work

3.3.2 Library construction and selection

The library used for the isolation of suppressor translocases was constructed using a modified procedure as described elsewhere (*71, 116*). Briefly, error-prone PCR was used to generate a *tatABC* library of ~5.5 x 10^6 members in the pTatABC vector. Following construction, the plasmid library was transformed into ElectroMAX DH5α (Invitrogen) for library propagation and storage. The plasmid library was midiprepped from this DH5α host and used to transform electrocompetent DADE (Δ*tatABCDE*) cells already harboring the pSALect–ssTorA–α_3B–TEM1 partner plasmid. Transformed cells were plated in serial dilutions and cultured overnight in 500 mL of LB medium containing 10 μg/mL tetracycline, 25 μg/mL chloramphenicol, and .2% glucose. The following day, plates were counted to ensure the diversity of the library was maintained, whereas, the overnight liquid culture was diluted to the appropriate cellular density for selections and plated on a series of ampicillin concentrations. Plates were incubated overnight at 30^{o}C, at which time individual clones were picked from the 200μg/mL plate and incubated in liquid culture.

3.3.3 Plasmid curing and isolation of mutant **tat** *vectors*

The pSALect–ssTorA–α_3B–TEM1 plasmid was cured from each clone by culturing the strain in the absence of chloramphenicol for several generations, streaking these cultures on LB-agar plates containing only tetracycline at an elevated temperature of 42^{o}C for 24 hours, and subsequent overnight liquid culture of single colonies at 37^{o}C. All strains cured of the pSALect vector were unable to grow in LB supplemented with chloramphenicol, but retained growth in LB containing tetracycline, indicating the loss of the pSALect vector and continued possession of the mutated pTatABC vector. Plasmid DNA isolated from these cured strains were

backtransformed into empty DADE (Δ*tatABCDE*) cells for all downstream experiments.

3.3.4 Spot plating and growth studies

Growth conditions and spot plating procedures used in this study are as described in Chapter 2.3.2, unless otherwise mentioned.

3.3.5 Subcellular fractionation and Western blot analysis

Similarly, the methods for protein analysis by subcellular fractionation and Western blotting that were used in this study are as described in Chapter 2.3.3, unless otherwise mentioned.

CHAPTER 4

CHARACTERIZATION OF CLASS I AND CLASS II MUTANT TRANSLOCASES AND VERIFICATION OF BROAD-ACTING SUPPRESSORS

4.0 Introduction

With the discovery of the Tat pathway occurring only recently, there is still much to be learned about how it functions. Several important independent studies have emerged (*81, 125, 126*) suggesting that the TatBC complex serves as the initial receptor binding site for precursor recognition. However, numerous shortcomings were evident in each of these previous works, including the fidelity of the genetic reporter, the functionality of isolated "suppressor" mutants, and an overall lack of confirmatory evidence that such suppressor translocases interacted with defective and wild-type signal peptides.

Similarly, there have been a number of biochemical studies demonstrating preliminary evidence of distinct interactions between preproteins and the Tat machinery; however, the reliability of these interactions is frequently called into question due to the artificial environment in which the experiments are invariably performed. The most promising *in vitro* studies have utilized a reconstituted Tat export system in the form of inverted inner membrane vesicles (*80, 127*). While these are certainly a step in the right direction, there remains to be a quality genetic means of directly assaying these interactions.

Previous attempts to genetically link quality control to the Tat machinery by isolating suppressor mutants has largely failed primarily due to the lack of a quality genetic screen or selection. For example, Kreutzenbeck *et al.* developed a genetic reporter system using the non-native Tat substrate ssTorA(KQ)-MBP and reported the

successful isolation of suppressor translocases with a relaxed signal peptide specificity (*126*). Despite these claims, their actual data suggested only a very limited ability to tolerate defective Tat signal peptides with merely one mutant having a marked handicap in managing this defective signal peptide – 44% export compared to a wild-type translocase and wild-type –RR– signal peptide. Moreover, in the development of their system, a wild-type translocase was fully able to export a defective twin-lysine signal peptide (*126*), suggesting that the screen was not as sensitive as they were claiming. Furthermore, this mutant was able to export about 25% more protein with a wild-type –RR– signal sequence, implying that the mutant may actually be hypersecreting all substrates and not actually relaxing quality control.

A second study created a randomized library of only the *tatC* gene and screened it for increased fluorescence on the FACS using the substrate ssTorA(KK)-GFP-SsrA, a substrate previously shown to result in a blockage of export, followed by immediate degradation in the cytoplasm by the ClpXP protein degradation system as a consequence of the SsrA tag (*125*). However, a major drawback of this work was the isolation of suppressors which had lost their ability to export native Tat substrates, suggesting that the directed evolution strategy that was used resulted in a shift in functionality rather than an expansion of functionality. Furthermore, similar to the work of Kreutzenbeck *et al.* (*126*), all of the suppressors reported in this work were only weak suppressors with at most a 20% improvement in function.

Thus, while attempts have been made to validate the early biochemical studies showing Tat-mediated substrate interaction, there remains to be the successful development of a high-quality genetic means of corroborating these early observations and a lack of direct genetic evidence. The work presented herein attempts to provide that missing element by thoroughly analyzing the best suppressors of folding quality control that were isolated from our genetic selection described in Chapters 2 and 3.

By using a family of proteins all derived from the same parental platform, all with comparable expression levels, we believe that our system exceeds previous attempts in that it is able to discern the export fitness of a target substrate based solely on the degree of folding character, offering the first direct genetic evidence of Tat-mediated folding quality control. Furthermore, while the best suppressors are able to export a number of previously transport-incompetent substrates, they otherwise retain all native functionality tested, as will be described in this chapter, indicating that they are indeed genuine suppressors of folding quality control.

4.1 Native functionality of class I suppressors

One of the major concerns in several of the previous studies on suppressor mutants was the inability of the isolated suppressors to transport native Tat substrates. It is well-known that *E. coli* cells lacking a functional Tat export pathway are unable to transport the N-acetylmuramoyl-L-alanine amidases, AmiA and AmiC, responsible for cleavage of the N-acetylmuramic acid in peptidoglycan (*125, 128-130*). This deficiency results in a cell's inability to properly divide during cell division and gives rise to a characteristic chain-like phenotype. As such, plasmid DNA from each mutant clone was separately purified and back-transformed into DADE (Δ*tatABCDE*) cells lacking any other plasmid in order to test for complementation of the chain phenotype. All mutants showed full complementation of the chain phenotype (Figure 4.1a), suggesting that each class I suppressor retained its native Tat functionality.

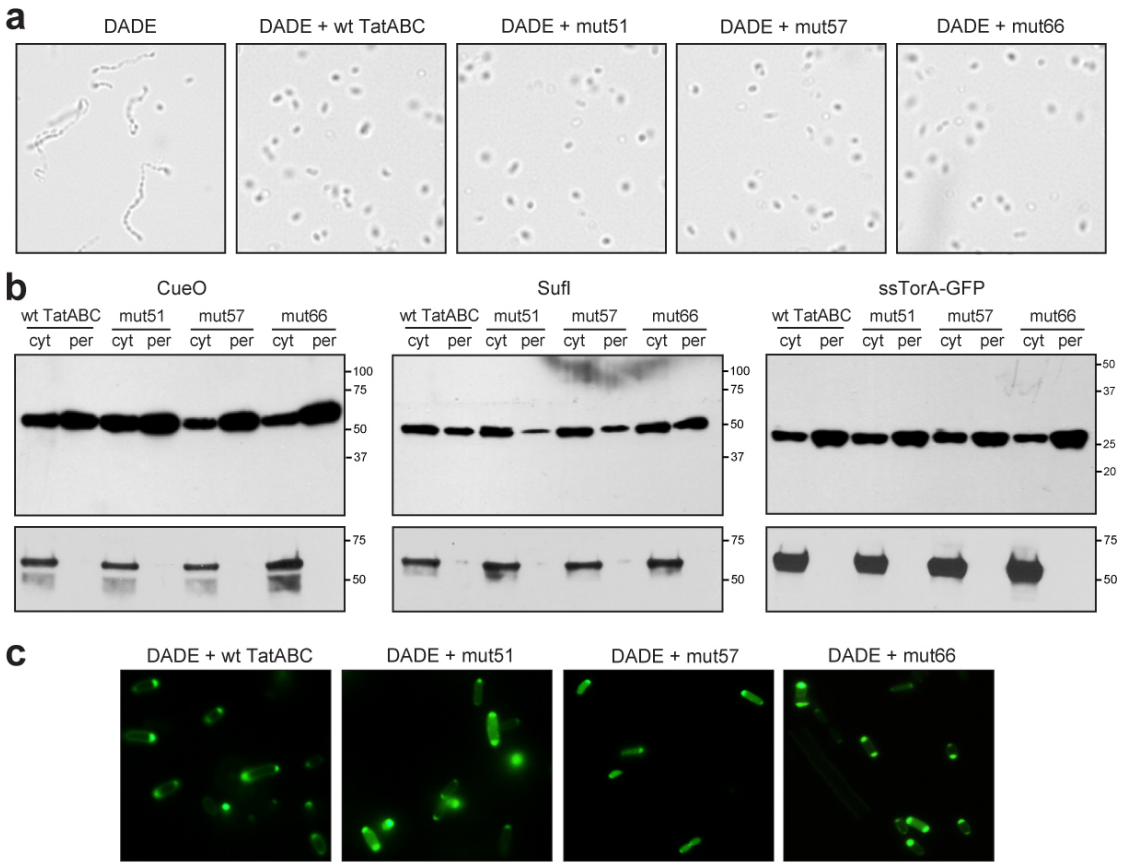

Figure 4.1 Class I suppressor export analysis of native Tat substrates. (a) Phase-contrast micrographs demonstrating the complementation of the chain phenotype characteristic of *tat* null strains. DADE (Δ*tatABCDE*) cells harboring functional Tat machinery show a wild-type singlet phenotype; whereas, empty DADE (Δ*tatABCDE*) cells have a characteristic chain phenotype due to the inability to export the Tat-specific amidases, AmiA and AmiC. All suppressor mutants complemented this chain phenotype, including class I and class II suppressors. (b) Western blot analysis of two native Tat substrates, CueO and SufI, along with the popular non-native substrate, ssTorA-GFP, revealed export tolerance among all class I suppressors. CueO and SufI were probed with mouse monoclonal 1ºα-Flag M2-HRP diluted 1:3,000 and GFP was probed with rabbit polyclonal 1ºα-GFP-HRP diluted 1:25,000. (c) Fluorescence microscopy of the ssTorA-GFP construct demonstrated polar accumulation of green fluorescence as characteristic of Tat-mediated GFP export.

In addition to the amidases, we interrogated two model Tat substrates, CueO (YacK) and SufI (FtsP). While both proteins belong to the multicopper oxidase superfamily, SufI lacks the incorporation of metal cofactors (*131*), whereas CueO is

typically found bound to four copper atoms (*132*). Despite their affiliation with the same multicopper oxidase superfamily, these two quintessential Tat proteins have two disparate functions. CueO is thought to maintain the copper levels within the cell and thus act as a resistance mechanism to copper toxicity (*133*). However, SufI's putative function was only recently hypothesized to involve a structural role at the septal ring during cellular division (*131*). Regardless of function, both substrates are exclusive to the Tat pathway and serve as vital components for interrogating native export functionality of Tat translocases. Therefore, each class I suppressor was co-transformed with native CueO and SufI expressed from a pBAD33 vector (*134*), induced for several hours, and fractionated via the ice-cold osmotic shock procedure in order to determine the subcellular localization of these model Tat substrates. Indeed, both CueO and SufI were found at comparable levels in the periplasm of each class I suppressor as compared to a wild-type Tat translocase, suggesting native export functionality was retained (Figure 4.1b).

4.2 Class I suppressors possess non-native export functionality

In an effort to explore the true extent of each class I suppressor's native export functionality, we began to interrogate non-native Tat substrates. Green fluorescence protein (GFP) has been previously used as a robust substrate for measuring green fluorescence in *E. coli* and many other hosts. Despite its popular use as a reporter protein, it was not until 2000 that Feilmeier *et al.* demonstrated that GFP is inactive in the periplasmic space. More specifically, GFP is completely inactive in the periplasm due to its inability to fold in the periplasm when exported via the general secretory pathway (*135*); however, it remains fully folded and active in the periplasm when exported through the twin-arginine translocation pathway (*136, 137*). Interestingly, ssTorA-GFP is also at a minimum predominantly, if not exclusively, exported via the

Tat pathway. Therefore, ssTorA-GFP serves as the foundation for interrogating the ability of class I suppressors to export non-native Tat substrates. Comparable to a wild-type translocase, each class I suppressor was found to export this non-native Tat fusion into the periplasm at a level of approximately 50% of the total soluble protein as evidenced by Western blotting (Figure 4.1b, rightmost panel). Furthermore, a necessary component of Tat-mediated export is to determine whether the exported protein is folded, and therefore, is functional. In agreement with earlier studies (*136, 138, 139*), fluorescence microscopy of each class I suppressor revealed halo-like or polar sequestration of green fluorescence, indicative of fully folded, and functional periplasmic GFP (Figure 4.1c).

The archetypal non-native Tat substrate, ssTorA-PhoA, has previously set the bar as the defining substrate for quality control. With the need for disulfide bond formation (*140*) and a native tertiary structure of a homodimer (*141, 142*), PhoA requires delicate assembly in a unique folding environment in order to achieve a native fold. Furthermore, by replacing its native Sec leader peptide with a Tat signal sequence, it was shown to become an exclusive Tat substrate (*38*). For this reason, the ssTorA-PhoA fusion was used in the seminal quality control study, whereby DeLisa and coworkers first posited the existence of an intrinsic quality control mechanism resident in the actual Tat machinery proteins. They demonstrated that the inability of Tat-mediated alkaline phosphotase export was likely a result of the highly reducing nature of the bacterial cytoplasm (*38*). It is well known that the thioredoxin and glutaredoxin pathways naturally maintain the bacterial cytoplasm in a highly reducing state – an environment which strongly disfavors the formation of disulfide bonds (*4*). Thus, such proteins, required to fold in the cytoplasm, would be unable to properly fold prior to export via the Tat pathway. Indeed, upon toggling the redox state of the

cytoplasm to an oxidizing environment, export of ssTorA-PhoA was achieved, confirming the aforementioned hypothesis.

Accordingly, the benchmark for class I suppressors was established as the ability to allow export of this delicate substrate. Remarkably, all three class I suppressors overcame the redox-induced lack of folding, allowing Tat-mediated translocation of the previously transport-incompetent reduced-PhoA as depicted by Western blot analysis of the subcellular fractions (Figure 4.2a). Furthermore, compared to a wild-type translocase, each class I suppressor also allowed a greater accumulation of the preprotein in the cytoplasm – a hallmark of transport-competent substrates. However, immunoblotting alone does not demonstrate that the alkaline phosphotase present in the periplasm is also functional, or whether it was exported in a inactive globular form. Consequently, these same DADE (Δ*tatABCDE*) cells were spread on plates supplemented with 0.1 mM isopropyl-β-D-thiogalactoside inducer (IPTG) and 150μg/mL 5-bromo-4-chloro-3-indolyl phosphate (BCIP) in order to test the functionality of the periplasmic PhoA. BCIP is a synthetic, chromogenic substrate of alkaline phosphotase which causes bacterial colonies to change from a white hue to a rich shade of blue in the presence of functionality active, periplasmic alkaline phosphotase and has been used to identify functional, Tat-mediated alkaline phosphotase activity (*117*). Indeed, all three class I suppressors not only showed export of ssTorA-PhoA, but also changed a blue colony color, confirming the presence of folded and functionality active periplasmic PhoA (Figure 4.2b). Likewise, in agreement with these previous studies, DADE (Δ*tatABCDE*) cells harboring a wild-type translocase remained white.

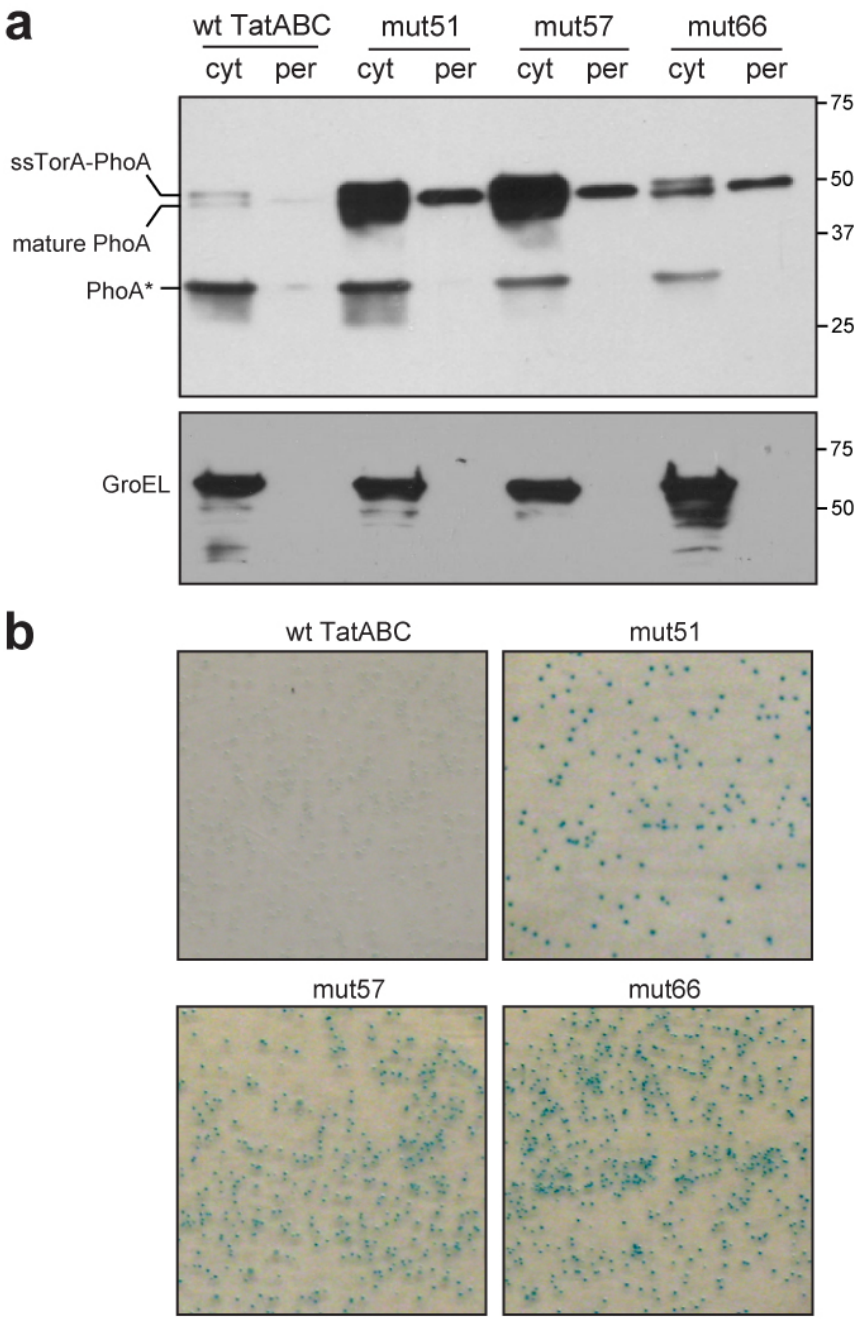

Figure 4.2 Alkaline phosphotase as a benchmark for class I suppressors. (a) DADE (Δ*tatABCDE*) cells harboring plasmid-encoded wild-type Tat machinery will not export Tat-directed PhoA as evidenced by Western blot. Conversely, Western blotting of class I suppressors reveals localization of ssTorA-PhoA into the periplasm. All samples were probed with rabbit polyclonal 1°α-PhoA-HRP diluted 1:5,000. (b) Furthermore, due to the export of folded and functional alkaline phosphotase, class I suppressors form blue colonies on LB-agar plates supplemented with appropriate antibiotics, 0.1 mM isopropyl-β-D-thiogalactoside (IPTG) and 150μg/mL 5-bromo-4-chloro-3-indolyl phosphate (BCIP).

4.3 Genetic dissection of class I suppressors reveals the underlying residues responsible for the suppressor phenotype

Plasmid DNA from the best ten mutants was isolated and sequenced to determine the location of each mutation in the *tatABC* gene cluster (Table 4.1). Most notably, 7 of 10 clones contained at least a single amino acid substitution at positions ranging from 94-99 of TatC, followed by at least one additional mutation in TatA and/or TatB. This area of clustering is of particular interest because it is in the first putative cytoplasmic loop of TatC (*52*), and thus, would very likely be an area to first interact with a substrate in order to judge folding character.

Table 4.1 Mutations present in the best ten suppressors.

Mutant	TatA	TatB	TatC	Class
Mut51	K82T	N119S	**P97S; A133T**	I
Mut57	K49M	**H109N; A122T; E147G; A162V**	**P8A; T62I; F94I; A98T; M122T; F136Y; N247S**	I
Mut66	**G2S; G3S**	K68M; E139K	I183T	I
Mut64	---	V12M; K30I; E102G	E103D	II
Mut65	Q75R; K77N	E120D; A153T	E15V; F94Y	II
Mut68	S35A; T78A	S7C	P97S; M159L; M163L	II
Mut63	Q68P; D85N; V89M	A38T; P128S; S136G	F94I; N247D	II
Mut78	I6V	T75M; N99D	L9I; L161Q; D248N	II
Mut73	K24R	A98S	F94L	II
Mut76	D46E	I36N; L39M; A43V A129V	A98T; S253C	II

Furthermore, this grouping of mutants consists of 4 mutations at F94, a site previously identified as essential for Tat export (*52, 125*). Due to the growing body of evidence suggesting TatC is a crucial component of this recognition and proofreading feature, our primary focus has been on TatC mutations, although, interesting mutations were eventually revealed elsewhere (Figure 4.3).

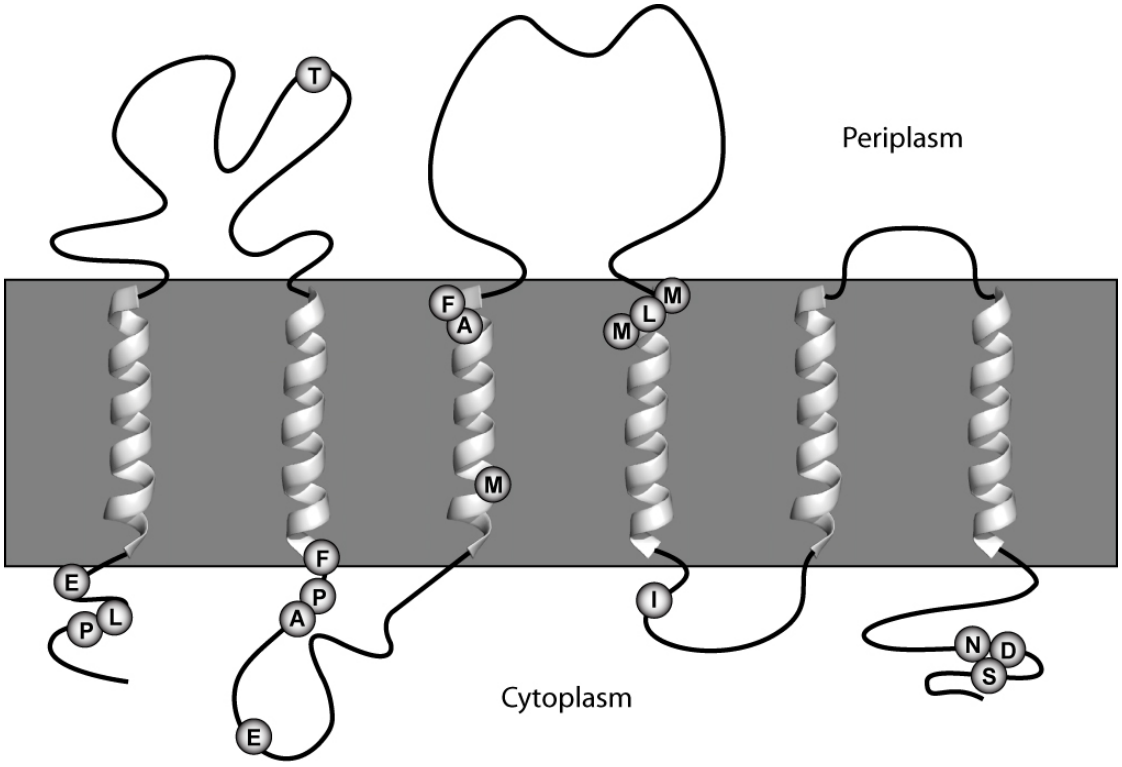

Figure 4.3 Schematic of TatC mutations from the best ten suppressors overlaid on a TatC topology map. The predicted topological structure of TatC as determined by Buchanan *et al.* (*52*), Behrendt *et al.* (*143*), and Ki *et al.* (*54*) was used to approximate the location of each TatC mutation pulled from the genetic selection. Based on these models, TatC is predicted to form six transmembrane α-helices with two cytoplasmic, and three periplasmic loops as depicted above. The bacterial cytoplasmic membrane is represented in the diagram by the uniformly shaded gray bar, with the cytoplasmic space shown below and the periplasmic space shown above. Strikingly, the majority of mutations are localized to the cytoplasmic domains, consistent with the hypothesis of Tat-mediated quality control.

Each class I suppressor was analyzed to determine the minimal suppressing residues (MSR) accountable for the class I export phenotype. Specifically, two restriction enzyme sites were introduced into the pTatABC plasmid by site-directed mutagenesis. With a restriction enzyme site now flanking each *tat* gene, the new vector, hereafter named pTatABC-XX, was used to systematically replace a mutated *tat* gene from each class I suppressor with its corresponding wild-type partner genes. For example, mutant 51 harbors two mutations in the *tatC* gene, namely P97S and A133T, along with N119S in *tatB* and K82T in *tatA*. Therefore, pTatABC* was created by PCR amplifying only the mutated *tatC* gene and pairing it with the wild-type versions of *tatA* and *tatB* to reveal a new variant harboring only the two *tatC* mutations, P97S and A133T. This same procedure was repeated to create mutants of each singly mutated *tat* gene and the two combinations, *tatAB* and *tatBC*, producing five new mutant translocases for each selected suppressor. All fifteen variants were co-transformed with the full panel of pSALect–ssTorA–α_3–TEM1 fusions and spot plated on varying concentrations of ampicillin in order to determine the MSR for each class I suppressor (Figure 4.4). As expected the MSR for both mutant 51 and 57 were found in the TatBC proteins, with mutant 51 allowing further simplification down to merely two residues within TatC. Strikingly, the MSR for mutant 66 was narrowed down to a pair of glycine-to-serine mutations (G2S and G3S) in TatA. These twin-serines are even more perplexing in that they are located in the periplasmic domain of TatA. Nevertheless, they are indeed the residues responsible for conferring suppression of the quality control feature for mutant 66.

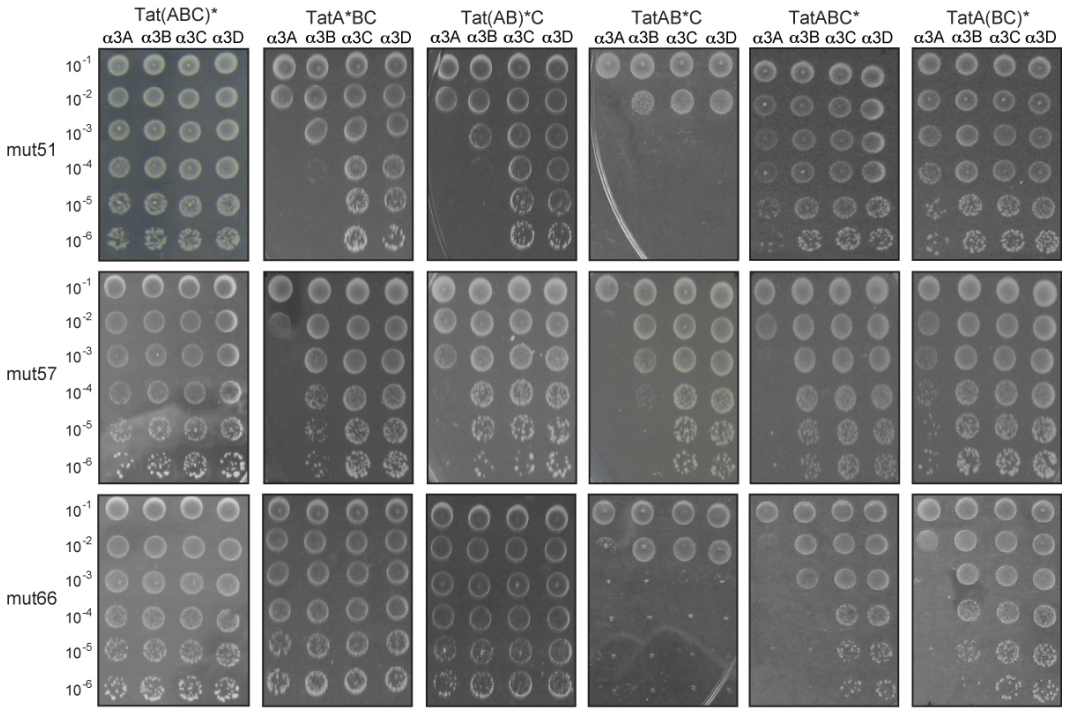

Figure 4.4 Genetic dissection of all class I suppressors at the single gene level. All cells in this study were a DADE (Δ*tatABCDE*) strain background harboring a mutant copy of the pTatABC plasmid and one of each of the pSALect–ssTorA–α_3–TEM1 plasmids, as indicated. Each class I suppressor was serial diluted by factors of ten, spotted onto 300 μg/mL ampicillin containing plates, and incubated overnight at 30ºC. Class I suppressors (leftmost panels), comprised of mutants 51, 57 and 66 were capable of transporting the least ordered, molten globular α_3A protein-fusion. As mentioned in the Materials and methods section, each selected translocase was subsequently genetically dissected by pairing a mutant gene with wild-type copies of the remaining *tat* counterparts. These new variants, *tatA*BC* (second column from the left), *tat(AB)*C* (third column from the left), *tatAB*C* (third column from the right), *tatABC** (second column from the right), or *tatA(BC)** (rightmost panels) were spotted in an identical manner as the class I selected mutants described above. The following dissected mutants were also capable of transporting the least ordered, molten globular α_3A protein-fusion, indicating that they were minimally necessary for suppression: pTatABC*-Mut51, pTatA(BC)*-Mut57, and pTatA*BC-Mut66. Genes marked with a * represent the mutant allele.

Interestingly, one of the best ten suppressors, mutant 73, was comprised of a single mutation in each of the three *tat* genes, allowing for a unique opportunity to dissect the contributions of not only each gene, but of each individual mutation (Figure 4.5). Ultimately, the MSR was identified as the contributions of not a single

mutation, but rather, the interaction of two mutations, one in TatB and one in TatC. This observation further validates the hypothesis that the intrinsic quality control feature of the Tat machinery resides within the TatBC multiprotein pre-complex. Remarkably, both single mutations (A98S in TatB and F94L in TatC) result in a loss of function, i.e., completely abolished export of all four α_3–TEM1 fusions. Thus, it is only the synergistic effect of both the A98S mutation in TatB and the F94L mutation in TatC that result in a reversal of function, allowing for suppression to occur. Once again, this underlines the benefit of using a directed co-evolution approach in order to facilitate the isolation of genuine suppressors that naturally function as a multiprotein complex.

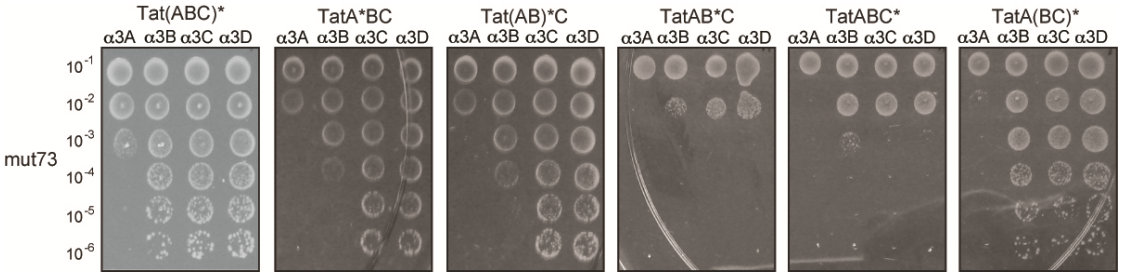

Figure 4.5 Genetic dissection of a class II suppressor at single amino acid substitution resolution. All cells in this study were a DADE (Δ*tatABCDE*) strain background harboring a mutant copy of the pTatABC plasmid and one of each of the pSALect–ssTorA–α_3–TEM1 plasmids, as indicated. Mutant 73, was identified as a class II suppressor by its ability to export only the α_3B protein-fusion and its well-ordered counterparts. Furthermore, this mutant was sequenced to reveal only a single amino acid substitution in each *tat* gene. Cells carrying the selected mutant 73 translocase were serial diluted by factors of ten, spotted onto 300 μg/mL ampicillin containing plates, and incubated overnight at 30ºC (leftmost panels). In a similar fashion as previously described, this translocase was subsequently genetically dissected by pairing a mutant gene with wild-type copies of the remaining *tat* counterparts. These new variants, *tatA*BC* (second column from the left), *tat(AB)*C* (third column from the left), *tatAB*C* (third column from the right), *tatABC** (second column from the right), or *tatA(BC)** (rightmost panels) were spotted in an identical manner as previously described. Only the combination of mutant *tatB* and *tatC* together, but not individually, were capable of transporting at least the α_3B protein-fusion. Genes marked with a * represent the mutant allele.

4.4 Discussion

In agreement with these earlier studies (*81, 125, 126*), suggesting that the TatBC complex serves as the initial receptor binding site for precursor recognition, the genetic selection described herein has enabled the identification of two such putative regions in TatB and TatC that may be responsible for judging the folding character of Tat substrates, namely, residues 94-99 of TatC and 100-130 of TatB. As both of these sites remain cytoplasmic and were heavily mutated among the best suppressors post-selection, they serve as a great starting point for further investigation. Consequently, it was entirely expected to find the MSR of two class I suppressors to reside within the TatBC proteins. Furthermore, both of these suppressors harbor mutations in the 94-99 region of TatC, an area previously regarded as essential for Tat-mediated export, validating the integrity of our genetic selection.

However, our genetic selection was also able to produce an absolutely unexpected class I suppressor, namely mutant 66, possessing its MSR in the periplasmic domain of TatA. While it remains undetermined, these twin-serines could interact with either the exported substrate at a later stage in translocation, i.e., after TatA has been recruited to the TatBC pre-complex, or the actual TatBC pre-complex itself. Although it initially seems counterintuitive to the current working model of Tat recognition and export, the former scenario remains an open possibility supported by recent evidence of an apparent retrograde translocation mechanism in the plant thylakoid Tat system. Specifically, using an *in vivo* tobacco protoplasts model, Di Cola and coworkers (*144*) demonstrated a late-stage, large-scale translocation reversal with both authentic and heterologous Tat substrates, whereby substrates were shown to be transported across the thylakoid membrane in a Tat-dependent manner, followed by a late-stage rejection after translocation had occurred and final localization back to the stromal side (since the Tat pathway in chloroplasts is a protein import system, the

stromal compartment is equivalent to the bacterial cytoplasmic space, and the chloroplast lumen is analogous to the bacterial periplasm). These studies provide substantial evidence that recruitment of the intact multi-protein Tat complex does not necessarily result in unidirectional translocation, and thus, allows for the possibility of a late-stage quality control event to occur. In other words, Tat proteins that are ultimately rejected by the translocation machinery have presumably interacted with a fully assembled TatABC translocon. If an analogous late-stage retrograde translocation process occurs in the *E. coli* Tat system on a similarly large-scale, TatA could serve as a late-stage quality control checkpoint such that key mutations on the periplasmic side of TatA (e.g., mutant 66) could allow for the suppression of this quality control event after TatABC assembly has already occurred. Nevertheless, the second alternative is much more straightforward in that conformational changes induced by the interaction of mutant TatA with TatB, TatC, or both could relax the intrinsic quality control mechanism of the TatBC components, allowing for the export of bound substrates. Regardless of the mechanism, this observation that mutant 66 possesses a MSR in TatA presents an amendment to the current Tat model of quality control and export. Namely, that the intrinsic quality control feature of the Tat machinery is perhaps not an "on-off sensor," appraising folding fidelity in a single step prior to export, but rather, a more complex series of "security checkpoints" first screened at the TatBC pre-complex, and perhaps minimally, again after TatABC assembly.

In summary, the α_3–TEM1 selection, developed in Chapter 3, has proven to be a robust tool for the isolation of mutant Tat translocases that retain their native export functionality while suppressing quality control. This was demonstrated by their ability to tolerate both wild-type substrates and non-native substrates – a feat surpassing many earlier studies seeking the isolation of suppressor translocases. Furthermore, we

have isolated several distinct classes of suppressors, each harboring unique mutations responsible for suppression of quality control. Insight gained from this selection has enabled the discovery of two precise regions on both TatB and TatC attributable to the strongest suppressor phenotype. While the precise molecular mechanism of action remains elusive, the genetic elucidation of several key steps underlying this intrinsic folding quality control mechanism holds great promise for future discovery and offers a broad-range of engineering and biotechnological applications.

4.5 Materials and methods

4.5.1 Bacterial strains, plasmids, and growth conditions

Bacterial strains, plasmids and growth conditions used in this study are described in Chapter 3.3.1, unless otherwise mentioned. In order to genetically dissect the suppressor mutants, the pTatABC template vector needed to be made modular. Enabling this feat was the addition of an in-frame XhoI restriction enzyme site flanking TatC at its 5'-end and a similar in-frame XbaI site at the 5'-end of TatB, resulting in the cloning vector, pTatABC-XX. Each mutated suppressor gene was subsequently PCR amplified and backcloned into the pTatABC-XX cloning vector using the appropriate cutsites. Below is a full list of all strains and plasmids used in this study. All plasmids generated in this study were confirmed by sequencing.

Table 4.2 Strains and plasmids used in this study.

Strain or Plasmid	Description	Reference
DADE	MC4100 derivative lacking all *tat* genes	Ref. 123
pCueO-Flag	Native *E. coli* Tat substrate, CueO, cloned into a pBAD33 backbone with a C-terminal Flag Tag; Cm^R	Ref. 134
pSufI-Flag	Native *E. coli* Tat substrate, SufI, cloned into a pBAD33 backbone with a C-terminal Flag Tag; Cm^R	Ref. 134
pTorA-GFP	pBAD33 backbone encoding GFP with a N-terminal TorA signal sequence; Cm^R	Ref. 137
pTorA-AP	pTrc99A backbone encoding Δss-PhoA with a N-terminal TorA signal sequence; Amp^R	Ref. 38
pTatABC	pBR322 backbone encoding the wild-type *tatABC* genes with its native promoter; Tc^R	This work
pTatABC-XX	As pTatABC with XbaI (5'-) and XhoI (3'-) cutsites flanking the *tatB* gene	This work
pTatABC-Mut51	As pTatABC with quality control suppressing mutations	This work
pTatABC-Mut57	As pTatABC with quality control suppressing mutations	This work
pTatABC-Mut66	As pTatABC with quality control suppressing mutations	This work
pTatABC-Mut73	As pTatABC with quality control suppressing mutations	This work
pTatA*BC-Mut	As pTatABC with wild-type *tatBC* genes, but a mutated *tatA* gene from selection	This work
pTatAB*C-Mut	As pTatABC with wild-type *tatAC* genes, but a mutated *tatB* gene from selection	This work
pTatABC*-Mut	As pTatABC with wild-type *tatAB* genes, but a mutated *tatC* gene from selection	This work
pTat(AB)*C-Mut	As pTatABC with a wild-type *tatC* gene, but mutated *tatAB* genes from selection	This work
pTatA(BC)*-Mut	As pTatABC with a wild-type *tatA* gene, but mutated *tatBC* genes from selection	This work

4.5.2 Spot plating and growth studies

Growth conditions and spot plating procedures used in this study are as described in Chapter 2.3.2, unless otherwise mentioned.

4.5.3 Subcellular fractionation and Western blot analysis

The methods for protein analysis by subcellular fractionation and Western blotting that were used in this study are as described in Chapter 2.3.3, with two exceptions. First, all substrates were prepared in an identical manner as described in Chapter 2.3.3, but fractionated after only ~2-4 hours at 30ºC. Second, protein detection in this study was determined with the following primary antibodies: rabbit polyclonal 1ºα-GroEL (Abcam) diluted 1:10,000, mouse monoclonal 1ºα-Flag M2-HRP (Sigma-Aldrich) diluted 1:3,000, rabbit polyclonal 1ºα-GFP-HRP (Abcam) diluted 1:25,000, and rabbit polyclonal 1ºα-PhoA-HRP (Abcam) diluted 1:5,000. The only secondary antibody used in this study was goat α-rabbit-HRP (Promega) diluted 1:2,500.

4.5.4 BCIP experiments

DADE (Δ*tatABCDE*) cells containing pTorA-AP (*38*) (i.e., plasmid-encoded alkaline phosphotase, *phoA*, targeted to the Tat pathway via an N-terminal TorA signal sequence) and a pTatABC vector (i.e., wild-type, mut 51, mut 57, or mut 66) were grown overnight at 37^{o}C in LB supplemented with 10 μg/ml tetracycline and 100 μg/ml ampicillin. Cells were subcultured the following day, plated at single colony resolution on LB-agar plates containing the appropriate antibiotics, 0.1 mM isopropyl-β-D-thiogalactoside (IPTG), and 150μg/mL 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and grown overnight for approximately 18 hours as described elsewhere (*117*).

4.5.5 Optical and fluorescence microscopy

All DADE (Δ*tatABCDE*) cells harboring a pTatABC plasmid were imaged using a Carl Zeiss Axioskop 40 optical microscope with a Zeiss 100x/1,30 Oil Plan-NEOFLUAR lens and a SPOT FLEX digital camera from Diagnostic Instruments, Inc. (Sterling Heights, MI). Microscope slides were prepared with five microliters of live bacterial cells from an overnight culture, sandwiched with a coverslip, and imaged using an oil-immersion technique at a magnification of 1000x. Fluorescence microscopy was conducted in a similar fashion using the identical instrument supplemented with an EXFO X-Cite Series 120 lamp (Mississauga, Ontario) and a EGFP emission filter.

CHAPTER 5

RATIONAL DESIGN OF A CLASS II TWIN-ARGININE TRANSLOCATION QUALITY CONTROL SUPPRESSOR

5.0 Introduction

Many studies have sought to find inactivating amino acid substitutions in the TatB protein of the twin-arginine translocation pathway in an effort to determine the nature of sidechains essential for Tat-mediated export. While very few, if any, single amino acid substitutions have been found, not everything regarding TatB function remains unidentified. Most of the information that is known regarding the mechanism in which TatB functions stems from either truncation analyses (*55*) or cysteine-scanning mutagenesis studies (*145*).

One early truncation study proposed that only the first 70 amino acids of the 171 aa protein were absolutely essential for function, as determined by a sequential C-terminal truncation analysis of 10 amino acids at a time (*55*). More specifically, the function of each TatB mutant was determined based on its periplasmic trimethylamine N-oxide (TMAO) reductase activity as compared to a wild-type TatB strain. However, while truncations up to 100 amino acids from the C-terminus displayed activity above background, only TatB proteins lacking the final 30 amino acids from the C-terminus behaved in a manner that was almost indistinguishable from wild-type. Thus, in the context of this study, the term "essential" is rather loosely defined. Regardless of the precise location in which TatB function was impaired, TatB proteins lacking a progressively larger chunk of the C-terminus directly correlated with a progressively diminishing TMAO reductase activity (*55*). Interestingly, the approximate location at which TatB was considered essential is immediately

downstream of the predicted α-helical transmembrane domain and the putative amphiphatic helical-like domain. Only the C-terminal portion of TatB, which is largely regarded to form a random coil, was deemed unimportant for function.

With a similar goal of fully characterizing the domains of TatB in order to better determine protein function, Maldonado and coworkers recently demonstrated the use of a two-hybrid system to determine that the so-called "membrane-extrinsic" domain of TatB, i.e., residues 22-100, was able to self-interact in a parallel fashion with similarly positioned TatB molecules (*146*). Further, even severe truncations leaving only the first 55 residues were shown to still stably interact with TatC in complex.

Taken together, the above studies suggest that the N-terminal 100 residues of TatB are likely regarded as essential only from a purely structural viewpoint. In particular, the N-terminus has demonstrated three separate structural roles: 1) the ability to self-interact with additional TatB molecules; 2) the integration into the cytoplasmic bilipid membrane by its predicted N-terminal α-helix; and 3) the association with TatC in the form of the TatBC complex. This in turn implies that the N-terminal domains could be responsible for the overall stability of the early TatBC complex, while the C-terminus could be involved in another aspect of export, such as quality control. The behavior of key residues in the C-terminus of TatB that could be involved in folding quality control will be investigated in this chapter.

5.1 Rational design of a class II suppressor

A major victory in any genetic selection is to not only isolate a desired protein, or protein complex, but to also glean information about the underlying basic science involved. To this end, we sought to rationally design a suppressor of Tat-mediated quality control based on our observations and analysis of the mutations present in the

class I suppressors. It became increasingly apparent that the majority of the mutations in *tatB* formed a cluster similar to that of *tatC*, albeit more sparsely distributed than those in *tatC*. In particular, the region from approximately residue 100 to 130, contained more than twice the number of mutations than any other region of comparable length, as shown below with the mutations from the best ten suppressors overlaid at their approximate location on a TatB topology map (Figure 5.1).

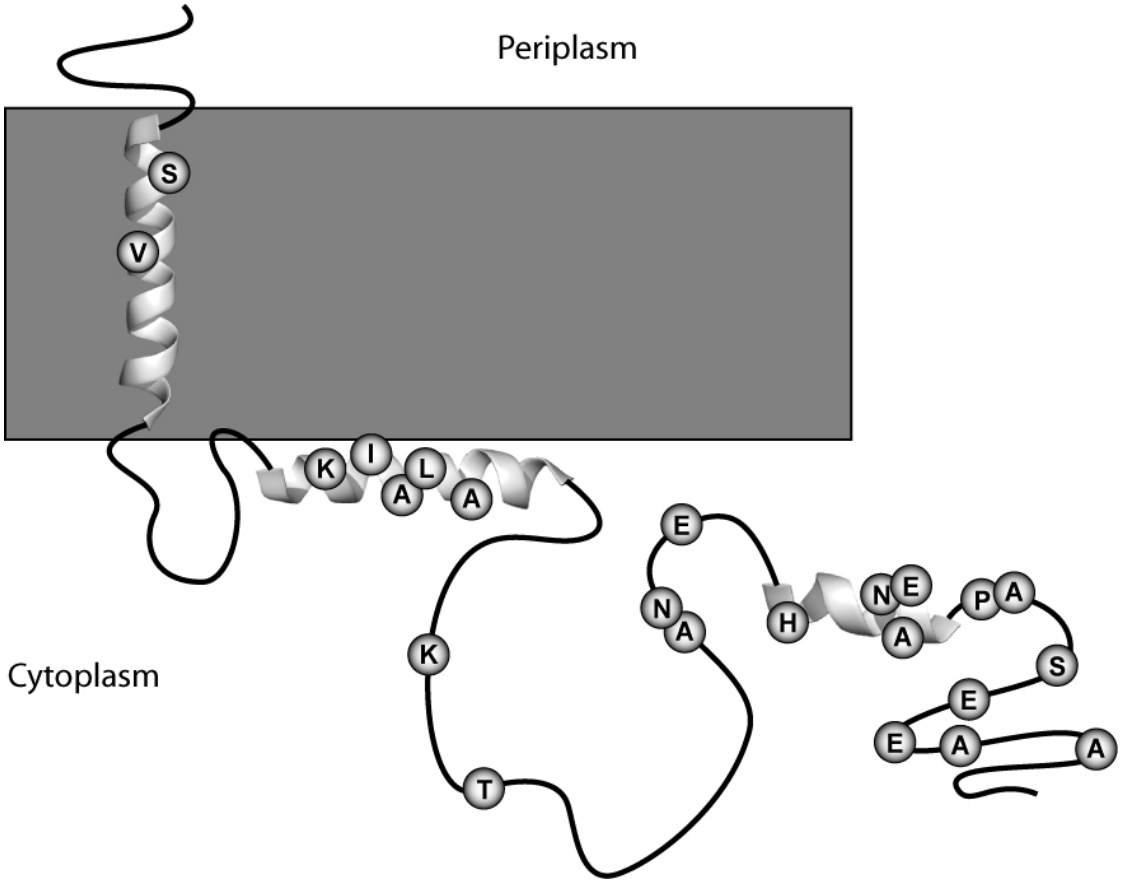

Figure 5.1 Schematic of TatB mutations from the best ten suppressors overlaid on a TatB topology map. The predicted topological structure of TatB as determined by computational predictions from Lee *et al*. (*55*) was used to approximate the location of each TatB mutation pulled from the genetic selection. As depicted above, the bacterial cytoplasmic membrane is represented by the uniformly shaded gray bar, with the cytoplasmic space shown below and the periplasmic space shown above. Consistent with the hypothesis of Tat-mediated quality control, the majority of mutations are localized to the cytoplasmic domain, and more specifically, to the C-terminal half of TatB.

In this region of TatB, there are three histidine residues that stood out for a variety of reasons. First, there are only seven histidines in the entire *tatABC* operon and three within a 14 amino acid stretch of the TatB protein alone. Second, there are an abundance of serine (or threonine) residues in this area of TatB, as well as in the cytoplasmic loops of TatC, and even more of these same mutations found in many of the class I suppressors. And perhaps most interesting is the observation that two of the histidines in TatB are spaced exactly 14 amino acids apart, each of which, are in turn flanked by a serine residue also spaced at a distance of 14 amino acids. Although the crystal structure remains unsolved, if this region were to at least form a transient alpha helix, it would form two serine-histidine pairs on the same side of a helical turn at exactly the *n* and *n+2* positions.

Stemming from these observations we made the following hypothesis: the imidazole side chains of the histidine residues found in the Tat proteins are, at a minimum, partly responsible for the strongest quality control suppressing phenotype. In order to test this hypothesis, a rationally designed, site-directed mutant was created by systematically replacing one, two and three of the TatB histidine residues with alanine residues (Figure 5.2a). Interestingly, with only a single histidine-to-alanine mutation, i.e., H123A, the mutant translocase was capable of suppressing the α_3B–TEM1 fusion and its more well-ordered counterparts. Furthermore, and most importantly, upon replacing all three histidine residues with alanines, the $TatB^{H109A/H113A/H123A}$ mutant translocase suppressed the full panel of α_3–TEM1 fusions up to an ampicillin concentration of 300 μg/mL (Figure 5.2b).

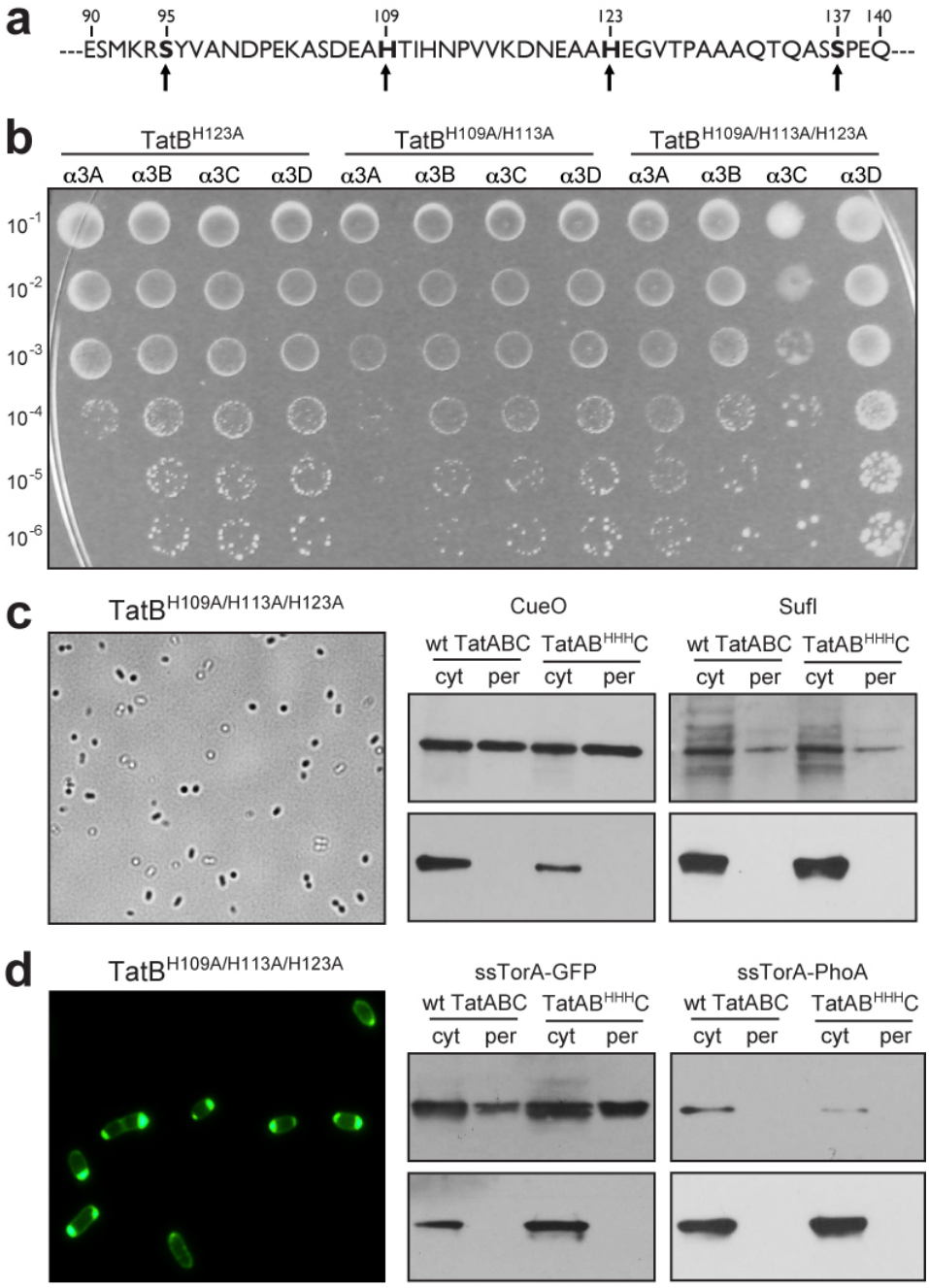

Figure 5.2 TatB histidine mutants result in class II suppression. (a) The *tatB* gene from residues 90 to 140, displaying the histidine and serine residues spaced exactly 14 amino acids apart. (b) Growth analysis of the site-directed H-to-A mutant translocases on LB-agar plates containing 300 μg/mL ampicillin. All cells are a DADE (Δ*tatABCDE*) strain background, harboring a pTatABC-Mutant translocase vector and a pSALect–ssTorA–α_3–TEM1 vector, as indicated. Five microliters of cells were spotted after ten-fold serial dilutions. (c & d) Microscopy and Western blot analysis of native and non-native substrates performed as described earlier.

In a similar fashion to the class I suppressors, the TatB$^{H109A/H113A/H123A}$ mutant translocase was characterized by its ability to export both native Tat substrates (AmiA, AmiC, CueO, and SufI), and non-native Tat substrates (ssTorA-GFP and ssTorA-PhoA). While the TatB$^{H109A/H113A/H123A}$ mutant translocase complemented the chain-like phenotype characteristic of DADE (Δ*tatABCDE*) cells and exported the native Tat substrates with equal facility as a wild-type translocase (Figure 5.2c), it was only partially permissive with the non-native Tat substrates (Figure 5.2d). In particular, ssTorA-GFP was very well tolerated and exported in a folded confirmation sequestered to the poles of the cell as evidenced by fluorescence microscopy. Moreover, immunoblotting of the subcellular fractions revealed ssTorA-GFP present in the periplasm at a level on par with the wild-type translocase. However, the defining feature of a class I suppressor is its ability to export ssTorA-PhoA in a folded and functional state and the TatB$^{H109A/H113A/H123A}$ mutant translocase failed this final test, resulting in the classification as a class II suppressor.

5.2 Discussion

In addition to the interesting location of many of these mutated residues, the identity of the actual residue is also quite remarkable. Histidines, serines, and aspartates are well known to interact with each other as part of a catalytic triad found in many serine proteases, such as trypsin, subtilisin, and acetylcholinesterase (*147*). However, it is unlikely that the abundance of serine (and threonine) mutations pulled from the genetic selection described in previous chapters, along with the rationally-designed histidine mutant described in this chapter, are acting in a catalytic manner to suppress quality control. Nevertheless, given their sheer abundance, the reactive hydroxyl sidechain chemistry of serine and threonine residues is likely a contributing factor in the overall suppressing phenotype.

Alternatively, histidines have also been shown to play essential roles in the activity of certain chaperones, such as calreticulin, a Ca^{+2}-binding chaperone found in the lumen of the endoplasmic reticulum (*148*). This latter example is particularly attractive due to the nature of the Tat machinery. In particular, as it was originally hypothesized in the seminal work of Tat-mediated quality control (*38*), a cytoplasmic domain of one or more Tat membrane proteins could function as a chaperone in its ability to identify folded substrates. Furthermore, Guo *et al.* suggest a possible mechanism of action for the essential histidine of calreticulin attributing its function as a chaperone to either the structural changes induced by the absence of the histidine residue or the inability of the histidine to bind zinc ions necessary for function (*148*). This is a particularly insightful observation because the majority of Tat substrates require the addition of a metal cofactor. Therefore, it is certainly plausible that the flexible cytoplasmic tail of TatB could act as a probe for metal cofactor addition enabled by the metal chelating ability of its histidine residues.

In this chapter, we have demonstrated the ability to rationally design a class II suppressor of folding quality control based on our observations gleaned from the best suppressors described in the previous chapters. This site-directed mutant, harboring three histidine mutations in the *tatB* gene, was able to weakly export the molten globular α_3 proteins as demonstrated by growth out to a dilution factor of 10^{-6} on 300μg/mL of ampicillin. This concentration of ampicillin was significantly higher than that of a wild-type translocase, but much lower than that of the best suppressors. Furthermore, the $TatB^{H109A/H113A/H123A}$ mutant translocase was capable of exporting the less cumbersome substrates, CueO, SufI and ssTorA-GFP, but remained unable to transport the archetypal class I substrate, ssTorA-PhoA and thus, was classified as a class II suppressor. Nevertheless, the ability to create such a mutant is an amazing feat and a testament to the genetic selection described in Chapter 3.

5.3 Materials and methods

5.3.1 Bacterial strains, plasmids, and growth conditions

Bacterial strains, plasmids and growth conditions used in this study are described in Chapter 3.3.1 and Chapter 4.5.1, unless otherwise mentioned.

5.3.2 Site-directed mutagenesis of **tatB**

The three histidine mutants were constructed using the QuikChange® II site-directed mutagenesis kit (Stratagene), whereby the native codons, CAC or CAT, encoding histidine residues were site-specifically mutated to the GCG codon, resulting in alanine residues. All mutants were generated in a wild-type pTatABC vector and sequenced for confirmation.

5.3.3 Spot plating and growth studies

Growth conditions and spot plating procedures used in this study are as described in Chapter 2.3.2, unless otherwise mentioned.

5.3.4 Subcellular fractionation and Western blot analysis

The methods for protein analysis by subcellular fractionation and Western blotting that were used in this study are as described in Chapter 2.3.3 and Chapter 4.4.3, unless otherwise mentioned.

5.3.5 Optical and fluorescence microscopy

All optical and fluorescence microscopy was performed exactly as described in Chapter 4.5.5, unless otherwise mentioned.

CHAPTER 6

POSSIBLE BIOTECHNOLOGICAL APPLICATIONS OF MUTANT TWIN-ARGININE TRANSLOCATION LIBRARIES

6.0 Introduction

Cost-effective biological hosts, such as gram negative bacteria, and especially *E. coli*, have long been used in the field of biotechnology to produce valuable recombinant proteins due to a variety of factors including: fast doubling times, ease of manipulation, high cellular densities, and simple protein purification techniques (*149*). One such class of valuable recombinant proteins that has garnered much interest in the biotechnology and pharmaceutical fields is the single-chain variable fragment antibody (scFv). These recombinant antibody molecules represent a substantial portion of proteins undergoing clinical trials (*150*). However, current production schemes generally involve purification from Chinese hamster ovary (CHO) cell lines, which are much more delicate and expensive to maintain than bacterial hosts. Despite these shortcomings, very large quantities of recombinant protein must be produced in order to be useful, and generally, this need far exceeds the ability of *E. coli* (*151*). Furthermore, scFv molecules normally require the formation of disulfide bonds in order to function, and thus, successful production of these proteins largely depends on exporting them into the periplasmic space (*152*). However, while export of these proteins provides a more conducive environment for folding, the quantity of protein exported is limited by the export process and is generally no more than 10% of the total soluble yield (*153*).

Therefore, many attempts have been made to improve the yield of recombinant antibody fragments in *E. coli*, e.g., refolding protein from inclusion bodies (*154*),

toggling the redox conditions of the cytoplasm (*38*), and re-engineering the protein fold to become stable within the cytoplasm (*155*), precluding disulphide bond formation. While many of these methods have achieved some degree of success, they generally require extensive work, and therefore, the cost for improving each individual substrate is greatly increased. Thus, a universal means for isolating large quantities of high purity recombinant scFvs by standard protein purification methods would serve as a vast improvement to the current methods and reduce the overall cost of producing many promising therapeutic targets that may otherwise never see the clinic.

To this end, we have screened a mutant *tatABC* library against a highly soluble scFv target in order to obtain broad-spectrum hypersecreting translocases that may help relieve the bottleneck associated with protein export and foster the discovery of novel therapeutics.

6.1 Isolation of hypersecreting Tat translocases that preserve quality control

While the isolation of suppressor translocases has proven to be invaluable in the identification of specific residues and particular regions of the Tat machinery proteins responsible for quality control at a basic science level, the true utility of the studies discussed in this body of work is the ability to exploit this knowledge to produce something of interest to the greater scientific community. To this end, we have developed the following hypothesis: if the Tat machinery is indeed an integral part of export quality control, it should be possible to isolate mutant translocases capable of hypersecreting valuable protein substrates while maintaining an intact quality control system. In other words, our next goal was to generate a completely new and unique set of mutant translocases that transport into the periplasm increased amounts of substrates possessing a common fold, instead of mutant translocases that suppress folding quality control.

To ensure that the diversity of the library was large enough to isolate meaningful clones, we took advantage of the identical *tatABC* library of ~5.5 x 10^6 members as described in Chapter 3. Previously, we used a particularly poorly folded substrate to select for mutant translocases that would necessitate the relaxation of the Tat machinery's intrinsic quality control mechanism in order for cells to remain viable on selective medium. Conversely, in an effort to avoid selection of suppressors, here we chose to use a very well-ordered protein substrate, facilitating the selection of hypersecreting variants. In particular, we chose to use the single chain variable fragment, scFv-R4. This antibody fragment was created by Martineau and coworkers (*152, 155*) through a rigorous four-round directed evolution strategy, culminating in the production of a very well-folded, highly soluble antibody fragment with great affinity for its target, β-galactosidase. Not only is this a very well-folded substrate, scFvs in general, are of similar size, shape, and fold. Thus, if our hypothesis held true, it is theoretically possible that each hypersecreting variant would be capable of exporting many different scFvs with similar faculty – a feat which holds immense potential for the pharmaceutical and biotechnology fields.

In a similar fashion as described in Chapter 3.1, prior to library selection cells harboring a wild-type Tat translocase and the pSALect–ssTorA–scFvR4–TEM1 partner plasmid were spotted onto a series of ampicillin concentrations in order to determine the appropriate selection conditions, i.e., a cellular dilution and ampicillin concentration that resulted in very little growth. This was determined to be 300μg/mL ampicillin and a cellular dilution factor of 10^{-3} from an overnight culture. Once these appropriate selection conditions were determined, DADE (Δ*tatABCDE*) cells containing the pSALect–ssTorA–scFvR4–TEM1 partner plasmid were transformed with the freshly midiprepped *tatABC* plasmid library and directly selected on ampicillin containing plates. Individual clones were picked and counter-screened at a

cellular dilution of 10^{-3} on LB-agar plates supplemented with 300μg/mL ampicillin, respectively, to rank-order and isolate only those clones which grew better than those cells harboring a wild-type Tat translocase. As expected, there was a variety of growth phenotypes, ranging from those that did not grow at these levels and those that grew rather well (Figure 6.1).

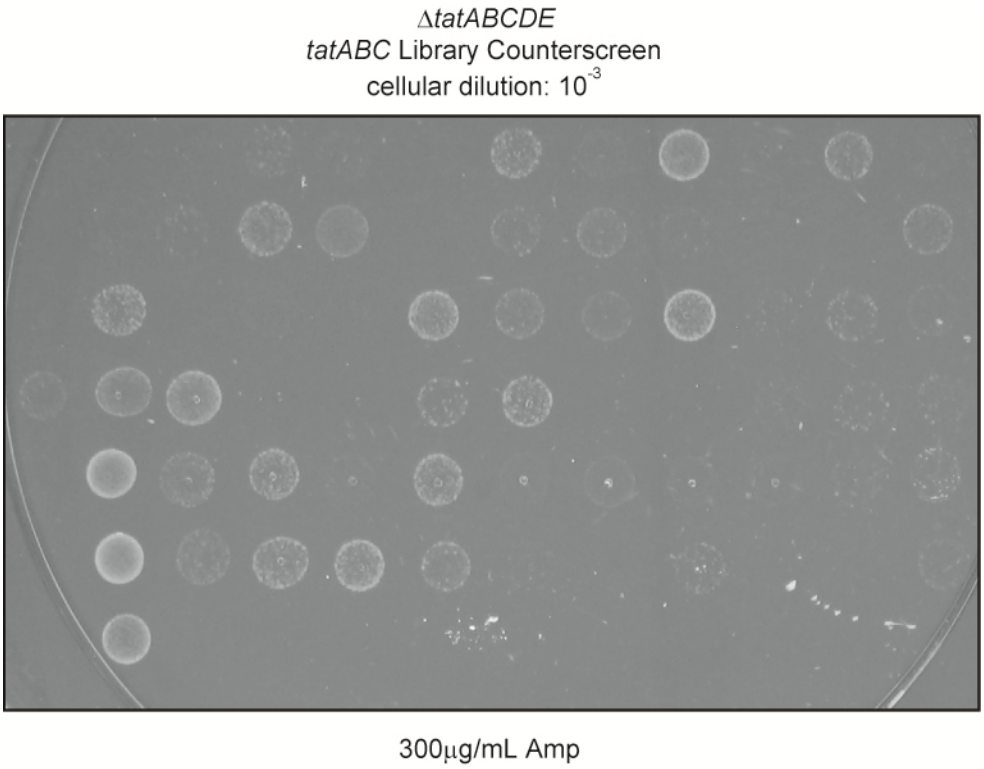

Figure 6.1 Counter-screen for hypersecreting variants reveals many potential candidates. Five microliters of cells from an overnight culture were diluted by a factor of 10^{-3}, spotted onto LB-agar plates supplemented with 300 μg/mL of ampicillin and growth overnight at 30°C.

The best three growers were isolated, cured of the pSALect–ssTorA–scFvR4–TEM1 partner plasmid, backtransformed into fresh DADE (Δ*tatABCDE*) cells harboring the pSALect–ssTorA–scFvR4–TEM1 vector to ensure that this improved growth phenotype was a result of the mutant translocase and not a mutation elsewhere

in the genome. Indeed, upon spotting these top three mutants on a series of ampicillin concentrations at every cellular dilution up to 10^{-8}, all mutant Tat translocases conferred a marked increase in resistance to ampicillin as determined by cellular growth in relation to a wild-type Tat translocase (Figure 6.2).

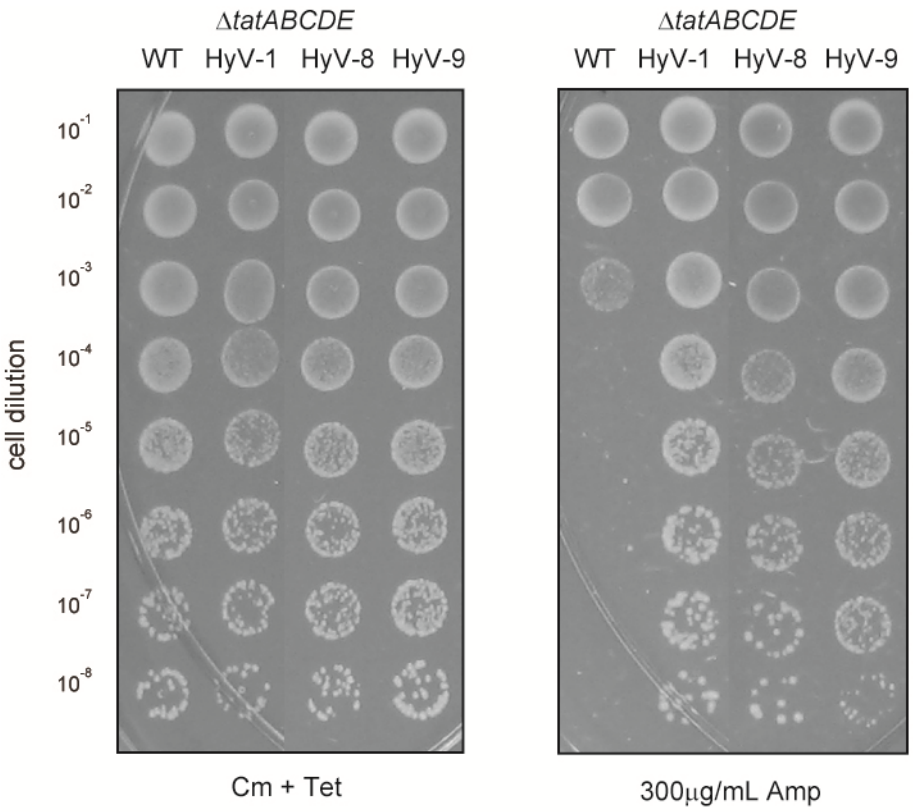

Figure 6.2 Demonstration of improved growth on ampicillin containing plates. Each hypersecreting variant was grown overnight at 37°C and serial diluted the following day by factors of ten in a 96-well tissue culture plate. Five microliters of cells were spotted onto LB-agar plates containing 25 μg/mL chloramphenicol and 10 μg/mL tetracycline (left) or 300μg/mL ampicillin (right) and grown overnight at 30°C.

These three hypersecreting variants were sequenced to determine the identity of the mutations responsible for this marked increased in growth on ampicillin (Table 6.1).

Table 6.1 Mutations present in the best hypersecreting variants.

Mutant	TatA	TatB	TatC
HyV-1	---	E8K; V32E; A122V; S167STOP	P85Q; K101N; A125T; I183T
HyV-8	L10M; D80Y	E53A	V128M; L218M; A251S
HyV-9	---	V1A; L9P; T72S; Q132H	S79T; R104H; L110M; L161P;

6.2 Discussion

Here we have shown a promising novel use of great engineering value for the genetic selection developed in Chapter 3 to investigate Tat translocases. Importantly, the system has been designed in a modular fashion in order to facilitate the selection of hypersecretors of any substrate protein, demonstrating its promise as a widespread tool for improving the folded, periplasmic yields of valuable target proteins.

A potential drawback of the technology described herein would be that in the process of creating a hypersecreting variant, one might also isolate a clone with a relaxed quality control mechanism, thereby eliminating the benefit of using the Tat pathway in the first place (i.e., exploiting its innate ability to export only fully folded, native proteins into the periplasm as a cellular pre-screen of folding quality). However, interestingly, the mutations we have observed thus far, suggest otherwise. In particular, the best three hypersecreting variants isolated in this proof-of-principle study all lack mutations in the 94-99 range of TatC and avoid mutation of the C-terminal half of TatB almost entirely. Moreover, there is only a single mutation in the hypersecreting lot that overlaps with those isolated as suppressors, namely, I183T in TatC. Interestingly, this mutation was also one of those carefully dissected in Chapter 4.3 and determined to have no effect on quality control as evidenced by a wild-type growth phenotype of the α_3s on ampicillin. Thus, at least in these preliminary

experiments, it is unlikely that Tat hypersecretors will also possess suppressor capabilities.

Another interesting observation among the hypersecretors is a dearth of cytoplasmic mutations, and instead, an abundance of hydrophobic residues within the transmembrane domains changed into methionine residues. A cursory glance would suggest that these are completely innocuous mutations; however, given the dynamics of protein export, and in particular, Tat export, these could be involved in a much more dominant way. The majority of these mutations are leucine-to-methionine or valine-to-methionine, and thus, are long, unbranched sidechain replacements. In other words, the ability of these residues to tightly pack within the transmembrane domains may be greatly disrupted, or minimally, alter the wild-type conformation within these regions. Moreover, it has been demonstrated that leucine-to-methionine substitutions can have a rather drastic destabilization effect on protein stability, especially in regions which lack solvent exposure (*156*), as would be the case with these transmembrane helices. Interestingly, the Tat export model gaining in popularity suggests that TatA homo-oligomers are responsible for destabilizing the membrane in the local proximity of the TatBC-substrate complex (*68*). Therefore, it is plausible that these methionine mutations found in transmembrane regions of the best hypersecretors result in the disruption of packing in the local micro-environment of Tat-substrate complexes, allowing for substrates to pass through the delicate membrane and into the periplasm. Further substantiating these claims are the recent observations by Lee and coworkers, demonstrating that the extreme N-terminus of TatB, as part of the TatBC complex, self-interacts within the membrane in such a way as to allow the possibility of membrane destabilizing mutations (*145*). Furthermore, the highly conserved glutamate residue at position 8 within this membrane domain has been implicated in linking proton movements to the transport process (*57, 157*), and

thus, certainly could be a candidate for hypersecreting activation. Most strikingly, two of the best hypersecreting variants isolated in this study have mutations in this area. Specifically, HyV-1, directly replaces this acidic glutamate with a basic lysine, which could have drastic effects on the hydrogen-bonding capabilities at this position within the transmembrane helix; and HyV-9, has a leucine-to-proline mutation one residue removed from Glu-8, which could have an equally membrane destabilizing effect due to the propensity of proline residues to form kinks in helices.

Taken together, the above results suggest that the hypersecreting variants indeed form a separate unique set of mutant Tat translocases responsible for hypersecreting their substrate protein, and not, merely suppressing quality control. Thus, by simply comparing the mutations found in any future hypersecreting variants with the list of mutant suppressors described in Chapter 4.3, one could eliminate any translocases which simultaneously hypersecrete substrate protein and suppress quality control. Moreover, the α_3 family of proteins could be used as a further screen for translocases, confirming that they continue to behave like wild-type in their ability to reject transport-incompetent substrates, while only hypersecreting fully folded substrates.

Therefore, in order to isolate hypersecretors it is recommended that one use a very well-ordered substrate to circumvent the need for suppression of the quality control mechanism. While quality control is likely to always be engaged, it is also likely that well-folded substrates spend less residence time at the machinery as compared to poorly folded substrates, and therefore, interact less with the machinery itself due the expeditious clearance through quality control. Conversely, poorly folded substrates are likely to dwell at the machinery, interact more with it, and ultimately fail to pass through quality control. Thus, while the strategy described herein remains a valid tool for isolating hypersecretors of any substrate protein, the best results will

likely occur when a well-ordered protein is used as the substrate. Furthermore, the most precious variants will likely be those hypersecreting valuable substrates of a common fold, such as the scFvs used in the proof-of-principle study described above, or cellulases – proteins heavily sought after for their potential use in generating renewable biofuels (*158*). Nonetheless, as demonstrated above the use of mutant *tatABC* libraries and robust genetic selections contained in this body of work hold great promise for many downstream biotechnological applications.

6.3 Materials and methods

6.3.1 Bacterial strains, plasmids, and growth conditions

Bacterial strains, plasmids and growth conditions used in this study are described in Chapter 3.3.1, unless otherwise mentioned. Below is a full list of all strains and plasmids used in this study. All plasmids generated in this study were confirmed by sequencing.

Table 6.2 Strains and plasmids used in this study.

Strain or Plasmid	Description	Reference
DADE	MC4100 derivative lacking all *tat* genes	Ref. 123
pSALect-scFvR4	ssTorA-scFvR4-TEM1 in pSALect; Cm^R	Refs. 71 & 152
pTatABC	pBR322 backbone encoding the wild-type *tatABC* genes with its native promoter; Tc^R	This work
pTatABC-HyV-1	As pTatABC with hypersecreting mutations as listed in Table 6.1	This work
pTatABC-HyV-8	As pTatABC with hypersecreting mutations as listed in Table 6.1	This work
pTatABC-HyV-9	As pTatABC with hypersecreting mutations as listed in Table 6.1	This work

6.3.2 Spot plating and growth studies

Growth conditions and spot plating procedures used in this study are as described in Chapter 2.3.2, unless otherwise mentioned.

6.3.3 Library construction and selection

The library used for the isolation of hypersecreting variants was the identical randomly mutated *tatABC* library as described in Chapter 3.3.2, unless otherwise mentioned.

6.3.4 Plasmid curing and isolation of mutant **tat** *vectors*

The pSALect–ssTorA–scFvR4–TEM1 plasmid was cured from each clone in order to isolate individual mutant *tatABC* plasmids as described in Chapter 3.3.3, unless otherwise mentioned.

CHAPTER 7

FUTURE DIRECTIONS OF TWIN-ARGININE TRANSLOCATION SUPPRESSORS OF FOLDING QUALITY CONTROL

7.0 Introduction

The identification of strong Tat suppressors of folding quality control, i.e., the class I suppressors described in detail in Chapter 4, holds great promise for the further elucidation of Tat-mediated export and intrinsic quality control. Within the previous chapters we have painstakingly characterized these mutant translocases in order to confirm their authenticity; however, despite the wealth of information that we have already gained from these initial experiments, there still remains more that can be learned from such transformative studies.

By comparison, over three decades ago elegant genetic strategies comparable to those described herein, were implemented to discover a series of protein localization mutants, termed *prl* suppressors (*159, 160*). These *prl* mutants were first isolated by their ability to export Sec substrates harboring defective signal peptides. Since little was known about the general secretory pathway at the time, these initial studies aimed to identify the gene products responsible for suppression through fine structure gene mapping. As might be expected, there were a variety of suppressors corresponding to three different genetic loci, termed *prlA* through *prlC*, with some suppressor phenotypes stronger than others (*159, 160*). Nevertheless, the ability to link genotype to phenotype provides strong evidence that the underlying molecular mechanisms were indeed a direct consequence of the genetic mutants found. Many studies later, the strongest of these Sec suppressors, *prlA*, was eventually determined to be a component of the inner-membrane general secretory apparatus, namely SecY

(*161-163*). In the context of our Tat suppressors of folding quality control, the true merit of these studies is that it took the isolation of Sec suppressor mutants to initiate the further elucidation of the molecular mechanics of Sec-mediated protein export (*164-167*).

Consequently, we would be remiss if we failed to acknowledge the considerable parallels our class I suppressors of the Tat pathway have with the *prl* mutants of the Sec pathway. With the isolation of these Tat suppressors of folding quality control, it is our sincere desire that further analysis of our most promising mutants will reveal important mechanistic details about Tat-mediated export in a similar fashion as the *prl* suppressor mutants have done for the Sec pathway. Several exciting avenues that can be immediately explored are outlined in the following experiments with the hope of sparking a new understanding of the export of folded proteins across the cytoplasmic membrane.

7.1 Specificity of quality control suppression

An important question asked early on in the selection process was: have we abolished all forms of quality control, such that every substrate is exported with little regard to specificity? The reason this becomes so important is that ultimately we sought to create a genetic selection which suppresses *folding* quality control, and thus, acts as a judge of the folding character alone. If every last substrate is exported regardless of the extent of misfolding, the genetic selection speaks very little to the underlying mechanism of suppression. To this end, we began probing substrate specificity by substituting defective signal peptides into the α_3 substrates using site-directed mutagenesis, whereby the characteristic twin-arginine residues were site-specifically mutated to either twin-lysines or a lysine-glutamine pair. It is well-documented that replacement of the near invariant twin-arginines with either twin-

lysines or a lysine-glutamine pair, completely blocks export via the Tat pathway. As such, regardless of the folding character of the α_3 protein, in theory, even a well-ordered variant harboring one of these defective signal peptides should not be exported, and thus, cells should not grow on ampicillin containing plates. Indeed, all class I suppressing translocases could not overcome either defective signal sequence as demonstrated by a substantially hindered growth pattern on ampicillin (Figure 7.1).

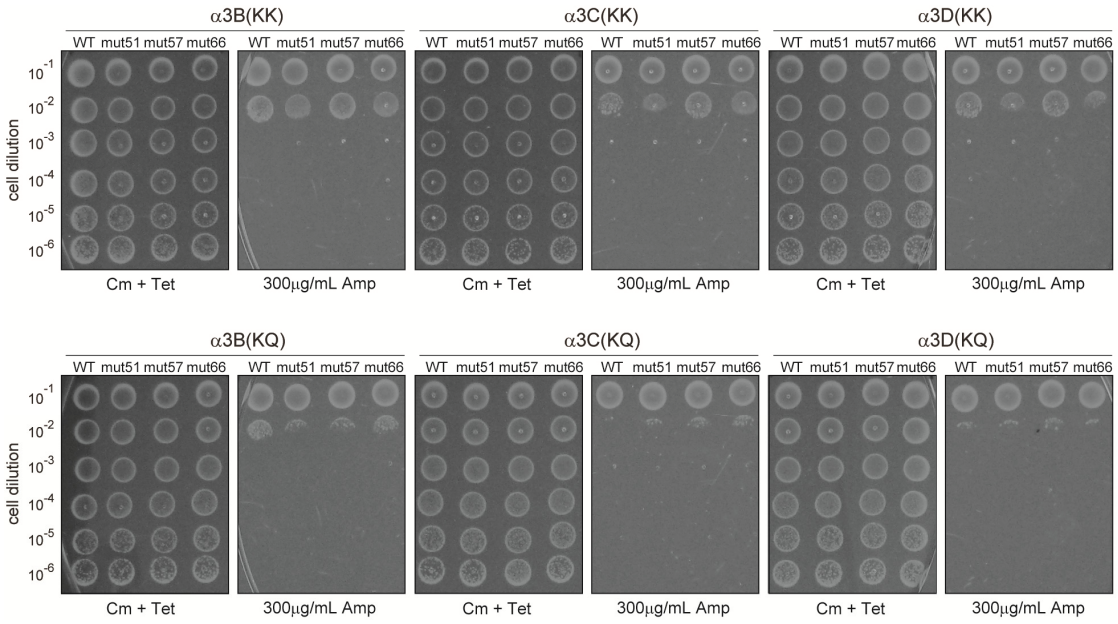

Figure 7.1 Class I suppressors retain the ability to recognize defective Tat signal sequences. All cells in this study were a DADE (Δ*tatABCDE*) strain background harboring a wild-type (WT), or mutant copy, of the pTatABC plasmid and one of the pSALect–ssTorA(KK/KQ)–α_3–TEM1 plasmids, as indicated. Each class I suppressor was serial diluted by factors of ten, spotted onto either 10 μg/mL tetracycline + 25 μg/mL chloramphenicol (positive growth control) or 300 μg/mL ampicillin containing plates, and incubated overnight at 30ºC. All translocases, including wild-type and each class I suppressor, lacked growth on ampicillin containing plates, indicative of an ability to recognize defective Tat signal peptides.

Retained ability to reject substrates lacking appropriate signal sequences confirms that each class I suppressor is truly judging folding character alone, and not, globally abolishing general quality control – further validation that the Tat machinery

proteins themselves indeed possess an intrinsic mechanism of folding quality control. In other words, it is likely that the suppressors described within this body of work are genuine suppressors of folding character and not completely relaxed bystanders in the translocation process such that they have become a passive channel for export. Nevertheless, an important follow-on to this question would be: what is the exact degree of substrate specificity achieved by these suppressors?

Apart from signal peptide recognition, another key element of Tat-mediated translocation is the export of cofactor-containing substrates. As previously mentioned, Tat substrates requiring proper cofactor insertion prior to export, such as NrfC and NapG, have been shown be to targeted for degradation when their cofactor clusters have been misassembled (*84*). Thus, a logical next question would be: can the class I suppressors isolated herein export these misassembled cofactor-containing Tat substrates? It is difficult to predict whether the class I suppressors would be able to accomplish this feat, as cofactor insertion is directly related to the protein fold; however, the degree of "misfolding" due to misassembly is not necessarily equivalent to "unfolded." Therefore, a series of experiments targeting the degree in which these suppressors are capable of probing cofactor assembly would help clarify the extent of suppression specificity. Moreover, if the class I suppressors are unable to export cofactor-misassembled substrates, the identical mutant *tatABC* library constructed in Chapter 3 could once again be used in order to isolate suppressors that allow export of these misassembled substrates.

7.2 Structural determinants of quality control suppression

While many of the studies described throughout the previous chapters have focused on determining a mechanism of action based on residue substitutions and the biochemical implications these may have on the substrate-machinery interaction, an

equally interesting notion pertains to interactions among machinery components themselves. Numerous studies have isolated Tat complexes of varying compositions with a molecular mass as large as approximately 600kDa in size, implying that the active translocon must contain multiple copies of each Tat protein (*66, 168, 169*). Furthermore, it has been demonstrated that the formation of a functional complex is intimately dependent upon a precise stoichiometry of the Tat proteins (*170-172*). Consequently, these studies have resulted in the development of robust methods for analyzing the molecular assembly and interactions between them. As such, a wealth of information could be gained from future experiments examining the perturbations that suppressors have on this stoichiometry and the implications these may have on the structure of complex formation. Furthermore, molecular visualization tools, such as the bimolecular fluorescence complementation (BiFC) technology developed to interrogate the molecular interactions along the Tat pathway (*173*), could be harnessed to study the specific contacts that each class of suppressor translocase makes with its substrates, as well as among the Tat components themselves. Structural determinants, whether they are based on differences in molar ratios among the Tat proteins in complex, interactions with each other in the context of substrate translocation, or something else altogether, could help further elucidate the precise molecular details of folding quality control, and more importantly, the suppression thereof.

7.3 Tat-specific proteolysis of misfolded substrates

To date, the research conducted herein has solely focused on protein folding quality control, that is, the productive process of creating a functional protein; however, as mentioned in Chapter 1, the inverse of protein folding, i.e., protein misfolding, also necessitates a form of quality control. The destructive process of protein degradation ultimately serves a constructive function to rid the cell of unusable

products and help alleviate the cellular stress associated with protein production, and thus, this proteolysis is a very important final form of protein folding quality control. Due to the uniqueness of the Tat pathway, the degradation of misfolded protein along the Tat pathway has the potential to be similarly special. In particular, if export of Tat substrates is mediated by an entire set of Tat-specific chaperones and machinery components, completely independent of those present in the Sec pathway, then it is a logical question to ask whether there also exists unique Tat-specific degradation machinery. One hypothesis recently suggested that TatD, a component of the Tat pathway long thought to serve little to no function in Tat-mediated export (*123*), may be involved in the rapid turnover of misfolded substrates to cellular proteases (*174*). However, shortly after publication the article was rescinded and its findings resolved to be the mere artifact of substrate overexpression levels as a consequence of inducer choice (*175*).

Nevertheless, the underlying question remains unanswered – what cellular components are responsible for the rapid degradation of misfolded, misassembled Tat substrates? Is it a Tat specific-protease, or a general cytoplasmic housekeeping mechanism? This intriguing aspect of folding quality control remains largely unexplored, and thus, studies aimed at shedding light on the black box that is Tat-mediated protein degradation of misfolded substrates promises to be a very rewarding endeavor.

7.4 Materials and methods

7.4.1 Bacterial strains, plasmids, and growth conditions

Bacterial strains, plasmids and growth conditions used in this study are described in Chapter 3.3.1 and Chapter 4.5.1, unless otherwise mentioned.

7.4.2 Site-directed mutagenesis of Tat signal sequences

The defective signal peptides used in this study were constructed using the QuikChange® II site-directed mutagenesis kit (Stratagene), whereby the native codons, CGT and CGG, encoding arginine residues were site-specifically mutated to the AAG AAG codons, or the AAG CAG codons, resulting in twin-lysines or a lysine-glutamine pair, respectively. All mutants were generated in a wild-type pTatABC vector and sequenced for confirmation.

7.4.3 Spot plating and growth studies

Growth conditions and spot plating procedures used in this study are as described in Chapter 2.3.2, unless otherwise mentioned.

REFERENCES

(1) Palmer, T., Sargent, F., and Berks, B. C. (2005) Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol 13*, 175-80.

(2) Rodrigue, A., Chanal, A., Beck, K., Muller, M., and Wu, L. F. (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J Biol Chem 274*, 13223-8.

(3) Berks, B. C., Sargent, F., and Palmer, T. (2000) The Tat protein export pathway. *Mol Microbiol 35*, 260-74.

(4) Ritz, D., and Beckwith, J. (2001) Roles of thiol-redox pathways in bacteria. *Annu Rev Microbiol 55*, 21-48.

(5) Dill, K. A., Ozkan, S. B., Weikl, T. R., Chodera, J. D., and Voelz, V. A. (2007) The protein folding problem: when will it be solved? *Curr Opin Struct Biol 17*, 342-6.

(6) Plaxco, K. W., Simons, K. T., and Baker, D. (1998) Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol 277*, 985-94.

(7) Xu, Y., Purkayastha, P., and Gai, F. (2006) Nanosecond folding dynamics of a three-stranded beta-sheet. *J Am Chem Soc 128*, 15836-42.

(8) Weikl, T. R., and Dill, K. A. (2007) Transition-states in protein folding kinetics: the structural interpretation of Phi values. *J Mol Biol 365*, 1578-86.

(9) Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999) Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature 400*, 693-6.

(10) Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature 431*, 590-6.

(11) Kramer, G., Rauch, T., Rist, W., Vorderwulbecke, S., Patzelt, H., Schulze-Specking, A., Ban, N., Deuerling, E., and Bukau, B. (2002) L23 protein functions as a chaperone docking site on the ribosome. *Nature 419*, 171-4.

(12) Hendrick, J. P., and Hartl, F. U. (1995) The role of molecular chaperones in protein folding. *Faseb J 9*, 1559-69.

(13) Vainberg, I. E., Lewis, S. A., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H. L., and Cowan, N. J. (1998) Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell 93*, 863-73.

(14) Hansen, W. J., Cowan, N. J., and Welch, W. J. (1999) Prefoldin-nascent chain complexes in the folding of cytoskeletal proteins. *J Cell Biol 145*, 265-77.

(15) Hartl, F. U., and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science 295*, 1852-8.

(16) Knoblauch, N. T., Rudiger, S., Schonfeld, H. J., Driessen, A. J., Schneider-Mergener, J., and Bukau, B. (1999) Substrate specificity of the SecB chaperone. *J Biol Chem 274*, 34219-25.

(17) Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997) Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *Embo J 16*, 1501-7.

(18) Wickner, S., Maurizi, M. R., and Gottesman, S. (1999) Posttranslational quality control: folding, refolding, and degrading proteins. *Science 286*, 1888-93.

(19) Pugsley, A. P. (1993) The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev 57*, 50-108.

(20) Hegde, R. S., and Bernstein, H. D. (2006) The surprising complexity of signal sequences. *Trends Biochem Sci 31*, 563-71.

(21) von Heijne, G. (1985) Signal sequences. The limits of variation. *J Mol Biol 184*, 99-105.

(22) Rapoport, T. A. (2007) Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature 450*, 663-9.

(23) Gelis, I., Bonvin, A. M., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A., and Kalodimos, C. G. (2007) Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell 131*, 756-69.

(24) Randall, L. L., and Hardy, S. J. (2002) SecB, one small chaperone in the complex milieu of the cell. *Cell Mol Life Sci 59*, 1617-23.

(25) Valent, Q. A., Scotti, P. A., High, S., de Gier, J. W., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B., and Luirink, J. (1998) The Escherichia coli SRP and SecB targeting pathways converge at the translocon. *Embo J 17*, 2504-12.

(26) Huber, D., Rajagopalan, N., Preissler, S., Rocco, M. A., Merz, F., Kramer, G., and Bukau, B. (2011) SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria. *Mol Cell 41*, 343-53.

(27) Van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004) X-ray structure of a protein-conducting channel. *Nature 427*, 36-44.

(28) Mitra, K., Frank, J., and Driessen, A. (2006) Co- and post-translational translocation through the protein-conducting channel: analogous mechanisms at work? *Nat Struct Mol Biol 13*, 957-64.

(29) Brundage, L., Hendrick, J. P., Schiebel, E., Driessen, A. J., and Wickner, W. (1990) The purified E. coli integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell 62*, 649-57.

(30) Erlandson, K. J., Miller, S. B., Nam, Y., Osborne, A. R., Zimmer, J., and Rapoport, T. A. (2008) A role for the two-helix finger of the SecA ATPase in protein translocation. *Nature 455*, 984-7.

(31) Schatz, G., and Dobberstein, B. (1996) Common principles of protein translocation across membranes. *Science 271*, 1519-26.

(32) Gu, S. Q., Peske, F., Wieden, H. J., Rodnina, M. V., and Wintermeyer, W. (2003) The signal recognition particle binds to protein L23 at the peptide exit of the Escherichia coli ribosome. *Rna 9*, 566-73.

(33) Huber, D., Boyd, D., Xia, Y., Olma, M. H., Gerstein, M., and Beckwith, J. (2005) Use of thioredoxin as a reporter to identify a subset of Escherichia coli signal sequences that promote signal recognition particle-dependent translocation. *J Bacteriol 187*, 2983-91.

(34) Lee, P. A., Tullman-Ercek, D., and Georgiou, G. (2006) The bacterial twin-arginine translocation pathway. *Annu Rev Microbiol 60*, 373-95.

(35) Berks, B. C. (1996) A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol 22*, 393-404.

(36) Berks, B. C., Sargent, F., De Leeuw, E., Hinsley, A. P., Stanley, N. R., Jack, R. L., Buchanan, G., and Palmer, T. (2000) A novel protein transport system involved in the biogenesis of bacterial electron transfer chains. *Biochim Biophys Acta 1459*, 325-30.

(37) Weiner, J. H., Bilous, P. T., Shaw, G. M., Lubitz, S. P., Frost, L., Thomas, G. H., Cole, J. A., and Turner, R. J. (1998) A novel and ubiquitous system for

membrane targeting and secretion of cofactor-containing proteins. *Cell 93*, 93-101.

(38) DeLisa, M. P., Tullman, D., and Georgiou, G. (2003) Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci U S A 100*, 6115-20.

(39) Hynds, P. J., Robinson, D., and Robinson, C. (1998) The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J Biol Chem 273*, 34868-74.

(40) Sanders, C., Wethkamp, N., and Lill, H. (2001) Transport of cytochrome c derivatives by the bacterial Tat protein translocation system. *Mol Microbiol 41*, 241-6.

(41) Fisher, A. C., and DeLisa, M. P. (2004) A little help from my friends: quality control of presecretory proteins in bacteria. *J Bacteriol 186*, 7467-73.

(42) Mould, R. M., Shackleton, J. B., and Robinson, C. (1991) Transport of proteins into chloroplasts. Requirements for the efficient import of two lumenal oxygen-evolving complex proteins into isolated thylakoids. *J Biol Chem 266*, 17286-9.

(43) Cline, K., Ettinger, W. F., and Theg, S. M. (1992) Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *J Biol Chem 267*, 2688-96.

(44) Ding, Z., and Christie, P. J. (2003) Agrobacterium tumefaciens twin-arginine-dependent translocation is important for virulence, flagellation, and chemotaxis but not type IV secretion. *J Bacteriol 185*, 760-71.

(45) Harrison, J. J., Ceri, H., Badry, E. A., Roper, N. J., Tomlin, K. L., and Turner, R. J. (2005) Effects of the twin-arginine translocase on the structure and antimicrobial susceptibility of Escherichia coli biofilms. *Can J Microbiol 51*, 671-83.

(46) Pradel, N., Ye, C., Livrelli, V., Xu, J., Joly, B., and Wu, L. F. (2003) Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic Escherichia coli O157:H7. *Infect Immun 71*, 4908-16.

(47) De Buck, E., Lebeau, I., Maes, L., Geukens, N., Meyen, E., Van Mellaert, L., Anne, J., and Lammertyn, E. (2004) A putative twin-arginine translocation pathway in Legionella pneumophila. *Biochem Biophys Res Commun 317*, 654-61.

(48) De Buck, E., Maes, L., Meyen, E., Van Mellaert, L., Geukens, N., Anne, J., and Lammertyn, E. (2005) Legionella pneumophila Philadelphia-1 tatB and tatC affect intracellular replication and biofilm formation. *Biochem Biophys Res Commun 331*, 1413-20.

(49) Ochsner, U. A., Snyder, A., Vasil, A. I., and Vasil, M. L. (2002) Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. *Proc Natl Acad Sci U S A 99*, 8312-7.

(50) Bronstein, P. A., Marrichi, M., Cartinhour, S., Schneider, D. J., and DeLisa, M. P. (2005) Identification of a twin-arginine translocation system in Pseudomonas syringae pv. tomato DC3000 and its contribution to pathogenicity and fitness. *J Bacteriol 187*, 8450-61.

(51) Allen, S. C., Barrett, C. M., Ray, N., and Robinson, C. (2002) Essential cytoplasmic domains in the Escherichia coli TatC protein. *J Biol Chem 277*, 10362-6.

(52) Buchanan, G., de Leeuw, E., Stanley, N. R., Wexler, M., Berks, B. C., Sargent, F., and Palmer, T. (2002) Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Mol Microbiol 43*, 1457-70.

(53) Drew, D., Sjostrand, D., Nilsson, J., Urbig, T., Chin, C. N., de Gier, J. W., and von Heijne, G. (2002) Rapid topology mapping of Escherichia coli inner-membrane proteins by prediction and PhoA/GFP fusion analysis. *Proc Natl Acad Sci U S A 99*, 2690-5.

(54) Ki, J. J., Kawarasaki, Y., Gam, J., Harvey, B. R., Iverson, B. L., and Georgiou, G. (2004) A periplasmic fluorescent reporter protein and its application in high-throughput membrane protein topology analysis. *J Mol Biol 341*, 901-9.

(55) Lee, P. A., Buchanan, G., Stanley, N. R., Berks, B. C., and Palmer, T. (2002) Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation. *J Bacteriol 184*, 5871-9.

(56) Settles, A. M., Yonetani, A., Baron, A., Bush, D. R., Cline, K., and Martienssen, R. (1997) Sec-independent protein translocation by the maize Hcf106 protein. *Science 278*, 1467-70.

(57) Sargent, F., Stanley, N. R., Berks, B. C., and Palmer, T. (1999) Sec-independent protein translocation in Escherichia coli. A distinct and pivotal role for the TatB protein. *J Biol Chem 274*, 36073-82.

(58) Bogsch, E. G., Sargent, F., Stanley, N. R., Berks, B. C., Robinson, C., and Palmer, T. (1998) An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem 273*, 18003-6.

(59) Sargent, F., Bogsch, E. G., Stanley, N. R., Wexler, M., Robinson, C., Berks, B. C., and Palmer, T. (1998) Overlapping functions of components of a bacterial Sec-independent protein export pathway. *Embo J 17*, 3640-50.

(60) Stanley, N. R., Findlay, K., Berks, B. C., and Palmer, T. (2001) Escherichia coli strains blocked in Tat-dependent protein export exhibit pleiotropic defects in the cell envelope. *J Bacteriol 183*, 139-44.

(61) Bruser, T., Yano, T., Brune, D. C., and Daldal, F. (2003) Membrane targeting of a folded and cofactor-containing protein. *Eur J Biochem 270*, 1211-21.

(62) Hou, B., Frielingsdorf, S., and Klosgen, R. B. (2006) Unassisted membrane insertion as the initial step in DeltapH/Tat-dependent protein transport. *J Mol Biol 355*, 957-67.

(63) Mori, H., and Cline, K. (2002) A twin arginine signal peptide and the pH gradient trigger reversible assembly of the thylakoid [Delta]pH/Tat translocase. *J Cell Biol 157*, 205-10.

(64) Dabney-Smith, C., Mori, H., and Cline, K. (2006) Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport. *J Biol Chem 281*, 5476-83.

(65) Gohlke, U., Pullan, L., McDevitt, C. A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H. R., and Berks, B. C. (2005) The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A 102*, 10482-6.

(66) Sargent, F., Gohlke, U., De Leeuw, E., Stanley, N. R., Palmer, T., Saibil, H. R., and Berks, B. C. (2001) Purified components of the Escherichia coli Tat protein transport system form a double-layered ring structure. *Eur J Biochem 268*, 3361-7.

(67) Bruser, T., and Sanders, C. (2003) An alternative model of the twin arginine translocation system. *Microbiol Res 158*, 7-17.

(68) Cline, K., and McCaffery, M. (2007) Evidence for a dynamic and transient pathway through the TAT protein transport machinery. *Embo J 26*, 3039-49.

(69) Alder, N. N., and Theg, S. M. (2003) Energetics of protein transport across biological membranes. a study of the thylakoid DeltapH-dependent/cpTat pathway. *Cell 112*, 231-42.

(70) McDonough, J. A., Hacker, K. E., Flores, A. R., Pavelka, M. S., Jr., and Braunstein, M. (2005) The twin-arginine translocation pathway of Mycobacterium smegmatis is functional and required for the export of mycobacterial beta-lactamases. *J Bacteriol 187*, 7667-79.

(71) Fisher, A. C., and DeLisa, M. P. (2009) Efficient isolation of soluble intracellular single-chain antibodies using the twin-arginine translocation machinery. *J Mol Biol 385*, 299-311.

(72) Fisher, A. C., Kim, W., and DeLisa, M. P. (2006) Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. *Protein Sci 15*, 449-58.

(73) Jack, R. L., Buchanan, G., Dubini, A., Hatzixanthis, K., Palmer, T., and Sargent, F. (2004) Coordinating assembly and export of complex bacterial proteins. *Embo J 23*, 3962-72.

(74) Hatzixanthis, K., Clarke, T. A., Oubrie, A., Richardson, D. J., Turner, R. J., and Sargent, F. (2005) Signal peptide-chaperone interactions on the twin-arginine protein transport pathway. *Proc Natl Acad Sci U S A 102*, 8460-5.

(75) Li, S. Y., Chang, B. Y., and Lin, S. C. (2006) Coexpression of TorD enhances the transport of GFP via the TAT pathway. *J Biotechnol 122*, 412-21.

(76) Maillard, J., Spronk, C. A., Buchanan, G., Lyall, V., Richardson, D. J., Palmer, T., Vuister, G. W., and Sargent, F. (2007) Structural diversity in twin-arginine signal peptide-binding proteins. *Proc Natl Acad Sci U S A 104*, 15641-6.

(77) Oresnik, I. J., Ladner, C. L., and Turner, R. J. (2001) Identification of a twin-arginine leader-binding protein. *Mol Microbiol 40*, 323-31.

(78) Holzapfel, E., Eisner, G., Alami, M., Barrett, C. M., Buchanan, G., Luke, I., Betton, J. M., Robinson, C., Palmer, T., Moser, M., and Muller, M. (2007) The entire N-terminal half of TatC is involved in twin-arginine precursor binding. *Biochemistry 46*, 2892-8.

(79) Richter, S., and Bruser, T. (2005) Targeting of unfolded PhoA to the TAT translocon of Escherichia coli. *J Biol Chem 280*, 42723-30.

(80) Panahandeh, S., Maurer, C., Moser, M., DeLisa, M. P., and Muller, M. (2008) Following the path of a twin-arginine precursor along the TatABC translocase of Escherichia coli. *J Biol Chem 283*, 33267-75.

(81) Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H. G., Brunner, J., and Muller, M. (2003) Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli. *Mol Cell 12*, 937-46.

(82) Cline, K., and Mori, H. (2001) Thylakoid DeltapH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *J Cell Biol 154*, 719-29.

(83) Gerard, F., and Cline, K. (2006) Efficient twin arginine translocation (Tat) pathway transport of a precursor protein covalently anchored to its initial cpTatC binding site. *J Biol Chem 281*, 6130-5.

(84) Matos, C. F., Robinson, C., and Di Cola, A. (2008) The Tat system proofreads FeS protein substrates and directly initiates the disposal of rejected molecules. *Embo J 27*, 2055-63.

(85) Hayes, S. A., and Dice, J. F. (1996) Roles of molecular chaperones in protein degradation. *J Cell Biol 132*, 255-8.

(86) Akiyama, Y., Yoshihisa, T., and Ito, K. (1995) FtsH, a membrane-bound ATPase, forms a complex in the cytoplasmic membrane of Escherichia coli. *J Biol Chem 270*, 23485-90.

(87) Danese, P. N., Snyder, W. B., Cosma, C. L., Davis, L. J., and Silhavy, T. J. (1995) The Cpx two-component signal transduction pathway of Escherichia coli regulates transcription of the gene specifying the stress-inducible periplasmic protease, DegP. *Genes Dev 9*, 387-98.

(88) Ruiz, N., and Silhavy, T. J. (2005) Sensing external stress: watchdogs of the Escherichia coli cell envelope. *Curr Opin Microbiol 8*, 122-6.

(89) Hochstrasser, M., and Wang, J. (2001) Unraveling the means to the end in ATP-dependent proteases. *Nat Struct Biol 8*, 294-6.

(90) Kim, J. Y., Fogarty, E. A., Lu, F. J., Zhu, H., Wheelock, G. D., Henderson, L. A., and DeLisa, M. P. (2005) Twin-arginine translocation of active human tissue plasminogen activator in Escherichia coli. *Appl Environ Microbiol 71*, 8451-9.

(91) Bryson, J. W., Desjarlais, J. R., Handel, T. M., and DeGrado, W. F. (1998) From coiled coils to small globular proteins: design of a native-like three-helix bundle. *Protein Sci 7*, 1404-14.

(92) Walsh, S. T., Cheng, H., Bryson, J. W., Roder, H., and DeGrado, W. F. (1999) Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proc Natl Acad Sci U S A 96*, 5486-91.

(93) Betz, S. F., Bryson, J. W., and DeGrado, W. F. (1995) Native-like and structurally characterized designed alpha-helical bundles. *Curr Opin Struct Biol 5*, 457-63.

(94) Beasley, J. R., and Hecht, M. H. (1997) Protein design: the choice of de novo sequences. *J Biol Chem 272*, 2031-4.

(95) Dahiyat, B. I., and Mayo, S. L. (1997) De novo protein design: fully automated sequence selection. *Science 278*, 82-7.

(96) Wei, Y., Liu, T., Sazinsky, S. L., Moffet, D. A., Pelczer, I., and Hecht, M. H. (2003) Stably folded de novo proteins from a designed combinatorial library. *Protein Sci 12*, 92-102.

(97) O'Neil, K. T., and DeGrado, W. F. (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*, 646-51.

(98) Hellinga, H. W. (1998) The construction of metal centers in proteins by rational design. *Fold Des 3*, R1-8.

(99) Regan, L. (1998) Proteins to order? *Structure 6*, 1-4.

(100) Choo, Y., Castellanos, A., Garcia-Hernandez, B., Sanchez-Garcia, I., and Klug, A. (1997) Promoter-specific activation of gene expression directed by bacteriophage-selected zinc fingers. *J Mol Biol 273*, 525-32.

(101) Pomerantz, J. L., Wolfe, S. A., and Pabo, C. O. (1998) Structure-based design of a dimeric zinc finger protein. *Biochemistry 37*, 965-70.

(102) Struthers, M. D., Cheng, R. P., and Imperiali, B. (1996) Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science 271*, 342-5.

(103) Harbury, P. B., Kim, P. S., and Alber, T. (1994) Crystal structure of an isoleucine-zipper trimer. *Nature 371*, 80-3.

(104) Fezoui, Y., Connolly, P. J., and Osterhout, J. J. (1997) Solution structure of alpha t alpha, a helical hairpin peptide of de novo design. *Protein Sci 6*, 1869-77.

(105) Kuroda, Y., Nakai, T., and Ohkubo, T. (1994) Solution structure of a de novo helical protein by 2D-NMR spectroscopy. *J Mol Biol 236*, 862-8.

(106) Kortemme, T., Ramirez-Alvarado, M., and Serrano, L. (1998) Design of a 20-amino acid, three-stranded beta-sheet protein. *Science 281*, 253-6.

(107) Gellman, S. H. (1998) Minimal model systems for beta sheet secondary structure in proteins. *Curr Opin Chem Biol 2*, 717-25.

(108) Lovejoy, B., Choe, S., Cascio, D., McRorie, D. K., DeGrado, W. F., and Eisenberg, D. (1993) Crystal structure of a synthetic triple-stranded alpha-helical bundle. *Science 259*, 1288-93.

(109) Harper, E. T., and Rose, G. D. (1993) Helix stop signals in proteins and peptides: the capping box. *Biochemistry 32*, 7605-9.

(110) Zhou, H. X., Lyu, P., Wemmer, D. E., and Kallenbach, N. R. (1994) Alpha helix capping in synthetic model peptides by reciprocal side chain-main chain interactions: evidence for an N terminal "capping box". *Proteins 18*, 1-7.

(111) Efimov, A. V. (1993) *Current Opinion in Structural Biology 3*, 379-384.

(112) Desjarlais, J. R., and Handel, T. M. (1995) De novo design of the hydrophobic cores of proteins. *Protein Sci 4*, 2006-18.

(113) Lutz, S., Fast, W., and Benkovic, S. J. (2002) A universal, vector-based system for nucleic acid reading-frame selection. *Protein Eng 15*, 1025-30.

(114) Santini, C. L., Ize, B., Chanal, A., Muller, M., Giordano, G., and Wu, L. F. (1998) A novel sec-independent periplasmic protein translocation pathway in Escherichia coli. *Embo J 17*, 101-12.

(115) Casadaban, M. J., and Cohen, S. N. (1979) Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: in vivo probe for transcriptional control sequences. *Proc Natl Acad Sci U S A 76*, 4530-3.

(116) Fisher, A. C., Rocco, M. A., and Delisa, M. P. (2010) Genetic selection of solubility-enhanced proteins using the twin-arginine translocation system. *Methods Mol Biol 705*, 53-67.

(117) Marrichi, M., Camacho, L., Russell, D. G., and DeLisa, M. P. (2008) Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria. *J Biol Chem 283*, 35223-35.

(118) Chen, G., Hayhurst, A., Thomas, J. G., Harvey, B. R., Iverson, B. L., and Georgiou, G. (2001) Isolation of high-affinity ligand-binding proteins by periplasmic expression with cytometric screening (PECS). *Nat Biotechnol 19*, 537-42.

(119) Berks, B. C., Palmer, T., and Sargent, F. (2005) Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Curr Opin Microbiol 8*, 174-81.

(120) Tarry, M. J., Schafer, E., Chen, S., Buchanan, G., Greene, N. P., Lea, S. M., Palmer, T., Saibil, H. R., and Berks, B. C. (2009) Structural analysis of substrate binding by the TatBC component of the twin-arginine protein transport system. *Proc Natl Acad Sci U S A 106*, 13284-9.

(121) Gerard, F., and Cline, K. (2007) The thylakoid proton gradient promotes an advanced stage of signal peptide binding deep within the Tat pathway receptor complex. *J Biol Chem 282*, 5263-72.

(122) Hakes, L., Lovell, S. C., Oliver, S. G., and Robertson, D. L. (2007) Specificity in protein interactions and its relationship with sequence diversity and coevolution. *Proc Natl Acad Sci U S A 104*, 7999-8004.

(123) Wexler, M., Sargent, F., Jack, R. L., Stanley, N. R., Bogsch, E. G., Robinson, C., Berks, B. C., and Palmer, T. (2000) TatD is a cytoplasmic protein with

DNase activity. No requirement for TatD family proteins in sec-independent protein export. *J Biol Chem 275*, 16717-22.

(124) Fromant, M., Blanquet, S., and Plateau, P. (1995) Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem 224*, 347-53.

(125) Strauch, E. M., and Georgiou, G. (2007) Escherichia coli tatC mutations that suppress defective twin-arginine transporter signal peptides. *J Mol Biol 374*, 283-91.

(126) Kreutzenbeck, P., Kroger, C., Lausberg, F., Blaudeck, N., Sprenger, G. A., and Freudl, R. (2007) Escherichia coli twin arginine (Tat) mutant translocases possessing relaxed signal peptide recognition specificities. *J Biol Chem 282*, 7903-11.

(127) Maurer, C., Panahandeh, S., Jungkamp, A. C., Moser, M., and Muller, M. (2010) TatB functions as an oligomeric binding site for folded Tat precursor proteins. *Mol Biol Cell 21*, 4151-61.

(128) Heidrich, C., Templin, M. F., Ursinus, A., Merdanovic, M., Berger, J., Schwarz, H., de Pedro, M. A., and Holtje, J. V. (2001) Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of Escherichia coli. *Mol Microbiol 41*, 167-78.

(129) Bernhardt, T. G., and de Boer, P. A. (2003) The Escherichia coli amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. *Mol Microbiol 48*, 1171-82.

(130) Priyadarshini, R., Popham, D. L., and Young, K. D. (2006) Daughter cell separation by penicillin-binding proteins and peptidoglycan amidases in Escherichia coli. *J Bacteriol 188*, 5345-55.

(131) Tarry, M., Arends, S. J., Roversi, P., Piette, E., Sargent, F., Berks, B. C., Weiss, D. S., and Lea, S. M. (2009) The Escherichia coli cell division protein and model Tat substrate SufI (FtsP) localizes to the septal ring and has a multicopper oxidase-like structure. *J Mol Biol 386*, 504-19.

(132) Roberts, S. A., Weichsel, A., Grass, G., Thakali, K., Hazzard, J. T., Tollin, G., Rensing, C., and Montfort, W. R. (2002) Crystal structure and electron transfer kinetics of CueO, a multicopper oxidase required for copper homeostasis in Escherichia coli. *Proc Natl Acad Sci U S A 99*, 2766-71.

(133) Singh, S. K., Grass, G., Rensing, C., and Montfort, W. R. (2004) Cuprous oxidase activity of CueO from Escherichia coli. *J Bacteriol 186*, 7815-7.

(134) Tullman-Ercek, D., DeLisa, M. P., Kawarasaki, Y., Iranpour, P., Ribnicky, B., Palmer, T., and Georgiou, G. (2007) Export pathway selectivity of Escherichia coli twin arginine translocation signal peptides. *J Biol Chem 282*, 8309-16.

(135) Feilmeier, B. J., Iseminger, G., Schroeder, D., Webber, H., and Phillips, G. J. (2000) Green fluorescent protein functions as a reporter for protein localization in Escherichia coli. *J Bacteriol 182*, 4068-76.

(136) Thomas, J. D., Daniel, R. A., Errington, J., and Robinson, C. (2001) Export of active green fluorescent protein to the periplasm by the twin-arginine translocase (Tat) pathway in Escherichia coli. *Mol Microbiol 39*, 47-53.

(137) DeLisa, M. P., Samuelson, P., Palmer, T., and Georgiou, G. (2002) Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem 277*, 29825-31.

(138) Santini, C. L., Bernadac, A., Zhang, M., Chanal, A., Ize, B., Blanco, C., and Wu, L. F. (2001) Translocation of jellyfish green fluorescent protein via the Tat system of Escherichia coli and change of its periplasmic localization in response to osmotic up-shock. *J Biol Chem 276*, 8159-64.

(139) Lee, Y.-F., Hsieh, H.-Y., Tullman-Ercek, D., Chiang, T.-K., Turner, R. J., Lin, S.-C. (2010) Enhanced translocation of recombinant proteins via the Tat pathway with chaperones in Escherichia coli. *Journal of the Taiwan Institute of Chemical Engineers 41*, 540-546.

(140) Sone, M., Kishigami, S., Yoshihisa, T., and Ito, K. (1997) Roles of disulfide bonds in bacterial alkaline phosphatase. *J Biol Chem 272*, 6174-8.

(141) Agrawal, D. K., and Wanner, B. L. (1990) A phoA structural gene mutation that conditionally affects formation of the enzyme bacterial alkaline phosphatase. *J Bacteriol 172*, 3180-90.

(142) Akiyama, Y., and Ito, K. (1993) Folding and assembly of bacterial alkaline phosphatase in vitro and in vivo. *J Biol Chem 268*, 8146-50.

(143) Behrendt, J., Standar, K., Lindenstrauss, U., and Bruser, T. (2004) Topological studies on the twin-arginine translocase component TatC. *FEMS Microbiol Lett 234*, 303-8.

(144) Di Cola, A., and Robinson, C. (2005) Large-scale translocation reversal within the thylakoid Tat system in vivo. *J Cell Biol 171*, 281-9.

(145) Lee, P. A., Orriss, G. L., Buchanan, G., Greene, N. P., Bond, P. J., Punginelli, C., Jack, R. L., Sansom, M. S., Berks, B. C., and Palmer, T. (2006) Cysteine-scanning mutagenesis and disulfide mapping studies of the conserved domain of the twin-arginine translocase TatB component. *J Biol Chem 281*, 34072-85.

(146) Maldonado, B., Kneuper, H., Buchanan, G., Hatzixanthis, K., Sargent, F., Berks, B. C., and Palmer, T. (2011) Characterisation of the membrane-extrinsic domain of the TatB component of the twin arginine protein translocase. *FEBS Lett 585*, 478-84.

(147) Gibney, G., Camp, S., Dionne, M., MacPhee-Quigley, K., and Taylor, P. (1990) Mutagenesis of essential functional residues in acetylcholinesterase. *Proc Natl Acad Sci U S A 87*, 7546-50.

(148) Guo, L., Groenendyk, J., Papp, S., Dabrowska, M., Knoblach, B., Kay, C., Parker, J. M., Opas, M., and Michalak, M. (2003) Identification of an N-domain histidine essential for chaperone function in calreticulin. *J Biol Chem 278*, 50645-53.

(149) Mansell, T. J., Fisher, A. C., and DeLisa, M. P. (2008) Engineering the protein folding landscape in gram-negative bacteria. *Curr Protein Pept Sci 9*, 138-49.

(150) Hudson, P. J. (1998) Recombinant antibody fragments. *Curr Opin Biotechnol 9*, 395-402.

(151) Ayala, M., Balint, R. F., Fernandez-de-Cossio, L., Canaan-Haden, J. W., Larrick, J. W., and Gavilondo, J. V. (1995) Variable region sequence modulates periplasmic export of a single-chain Fv antibody fragment in Escherichia coli. *Biotechniques 18*, 832, 835-8, 840-2.

(152) Martineau, P., and Betton, J. M. (1999) In vitro folding and thermodynamic stability of an antibody fragment selected in vivo for high expression levels in Escherichia coli cytoplasm. *J Mol Biol 292*, 921-9.

(153) Carter, P., Kelley, R. F., Rodrigues, M. L., Snedecor, B., Covarrubias, M., Velligan, M. D., Wong, W. L., Rowland, A. M., Kotts, C. E., Carver, M. E., and et al. (1992) High level Escherichia coli expression and production of a bivalent humanized antibody fragment. *Biotechnology (N Y) 10*, 163-7.

(154) Misawa, S., and Kumagai, I. (1999) Refolding of therapeutic proteins produced in Escherichia coli as inclusion bodies. *Biopolymers 51*, 297-307.

(155) Martineau, P., Jones, P., and Winter, G. (1998) Expression of an antibody fragment at high levels in the bacterial cytoplasm. *J Mol Biol 280*, 117-27.

(156) Gassner, N. C., Baase, W. A., and Matthews, B. W. (1996) A test of the "jigsaw puzzle" model for protein folding by multiple methionine substitutions within the core of T4 lysozyme. *Proc Natl Acad Sci U S A 93*, 12155-8.

(157) Mori, H., and Cline, K. (2001) Post-translational protein translocation into thylakoids by the Sec and DeltapH-dependent pathways. *Biochim Biophys Acta 1541*, 80-90.

(158) Maki, M., Leung, K. T., and Qin, W. (2009) The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. *Int J Biol Sci 5*, 500-16.

(159) Emr, S. D., Hanley-Way, S., and Silhavy, T. J. (1981) Suppressor mutations that restore export of a protein with a defective signal sequence. *Cell 23*, 79-88.

(160) Emr, S. D., and Bassford, P. J., Jr. (1982) Localization and processing of outer membrane and periplasmic proteins in Escherichia coli strains harboring export-specific suppressor mutations. *J Biol Chem 257*, 5852-60.

(161) Bieker, K. L., Phillips, G. J., and Silhavy, T. J. (1990) The sec and prl genes of Escherichia coli. *J Bioenerg Biomembr 22*, 291-310.

(162) Schatz, P. J., and Beckwith, J. (1990) Genetic analysis of protein export in Escherichia coli. *Annu Rev Genet 24*, 215-48.

(163) Ito, K., Wittekind, M., Nomura, M., Shiba, K., Yura, T., Miura, A., and Nashimoto, H. (1983) A temperature-sensitive mutant of E. coli exhibiting slow processing of exported proteins. *Cell 32*, 789-97.

(164) Sako, T. (1991) Novel prlA alleles defective in supporting staphylokinase processing in Escherichia coli. *J Bacteriol 173*, 2289-96.

(165) Derman, A. I., Puziss, J. W., Bassford, P. J., Jr., and Beckwith, J. (1993) A signal sequence is not required for protein export in prlA mutants of Escherichia coli. *Embo J 12*, 879-88.

(166) Brickman, E. R., Oliver, D. B., Garwin, J. L., Kumamoto, C., and Beckwith, J. (1984) The use of extragenic suppressors to define genes involved in protein export in Escherichia coli. *Mol Gen Genet 196*, 24-7.

(167) Akiyama, Y., and Ito, K. (1987) Topology analysis of the SecY protein, an integral membrane protein involved in protein export in Escherichia coli. *Embo J 6*, 3465-70.

(168) de Leeuw, E., Granjon, T., Porcelli, I., Alami, M., Carr, S. B., Muller, M., Sargent, F., Palmer, T., and Berks, B. C. (2002) Oligomeric properties and signal peptide binding by Escherichia coli Tat protein transport complexes. *J Mol Biol 322*, 1135-46.

(169) Bolhuis, A., Bogsch, E. G., and Robinson, C. (2000) Subunit interactions in the twin-arginine translocase complex of Escherichia coli. *FEBS Lett 472*, 88-92.

(170) Bolhuis, A., Mathers, J. E., Thomas, J. D., Barrett, C. M., and Robinson, C. (2001) TatB and TatC form a functional and structural unit of the twin-arginine translocase from Escherichia coli. *J Biol Chem 276*, 20213-9.

(171) De Leeuw, E., Porcelli, I., Sargent, F., Palmer, T., and Berks, B. C. (2001) Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway. *FEBS Lett 506*, 143-8.

(172) Oates, J., Mathers, J., Mangels, D., Kuhlbrandt, W., Robinson, C., and Model, K. (2003) Consensus structural features of purified bacterial TatABC complexes. *J Mol Biol 330*, 277-86.

(173) Kostecki, J. S., Li, H., Turner, R. J., and DeLisa, M. P. (2010) Visualizing interactions along the Escherichia coli twin-arginine translocation pathway using protein fragment complementation. *PLoS One 5*, e9225.

(174) Matos, C. F., Di Cola, A., and Robinson, C. (2009) TatD is a central component of a Tat translocon-initiated quality control system for exported FeS proteins in Escherichia coli. *EMBO Rep 10*, 474-9.

(175) Lindenstrauss, U., Matos, C. F., Graubner, W., Robinson, C., and Bruser, T. (2010) Malfolded recombinant Tat substrates are Tat-independently degraded in Escherichia coli. *FEBS Lett 584*, 3644-8.